SC552126

Mathematical Methods in Physics II

Physics KKU

Chakrit Pongkitivanichkul

Chapter 1: Vector Space

เราได้ศึกษาเกี่ยวกับเวกเตอร์และแคลคูลัสของเวกเตอร์ไปแล้วในวิชาก่อนหน้านี้ ในวิชานี้เราจะขยายแนวคิดของเวกเตอร์ไปสู่ขอบเขตที่เป็นนามธรรมมากขึ้น

1.1) ปริภูมิเวกเตอร์ (Vector Space)

ในทางคณิตศาสตร์เรามักจะพิจารณาศึกษาสมบัติที่คล้ายคลึงกันของวัตถุแต่ละประเภทเพื่อทำการจัดหมวดหมู่ เพื่อที่จะลดรูปโครงสร้างของเวกเตอร์ที่เรารู้จักเป็นรูปทั่วไปเพื่อจะปรับใช้กับวัตถุอื่นๆ เราจะเริ่มจากของสองอย่างที่จำเป็นในการมีเวกเตอร์คือ

- เวกเตอร์ (Vectors) (เรามักเขียนแทนด้วย $|v\rangle$ หรือ $\vec{v}$)
- ตัวเลข (Scalars)

(ระวังว่าตัวเลขก็ใช้แทนเวกเตอร์ได้ แต่การมีเวกเตอร์ต้องมีตัวเลขมาคู่กับมันเสมอ)

วัตถุที่เป็นเวกเตอร์มีความหลากหลายมากซึ่งอาจจะเป็นเวกเตอร์แบบลูกศรที่เราเคยรู้จัก อาจจะเป็นแมทริกซ์, เป็นพหุนามหรือแม้แต่เป็นฟังก์ชั่นก็ได้

ในขณะที่ตัวเลขหรือสเกล่าร์นั้นมีได้ไม่กี่รูปแบบเช่น จำนวนจริง (Real numbers) จำนวนเชิงซ้อน (Complex numbers) เป็นต้น โดยปรกติแล้วเรามักเรียกชนิดของวัตถุที่มีสมบัติเหมือนตัวเลขว่าฟีลด์ (Fields) ในที่นี้ความหมายของฟีลด์ไม่เหมือนสนามที่เรามักคุ้นเคยในฟิสิกส์

นอกเหนือจากของสองอย่างที่จำเป็นต่อการเป็นเวกเตอร์แล้ว เรายังต้องมีตัวดำเนินการ (operator) ที่เชื่อมโยงของสองอย่างเข้าหากัน (มักเรียกว่า binary operator) ซึ่งได้แก่

- การบวกเวกเตอร์ (Vector addition) $|a\rangle + |b\rangle = |c\rangle$
- การคูณเวกเตอร์ด้วยตัวเลข (Scalar multiplication) $\lambda\,|a\rangle = |c\rangle$

โดย binary หมายความว่ามีเวกเตอร์สองตัวมากระทำกันแล้วได้เวกเตอร์อีกตัว หรือ ตัวเลขมาคูณกับเวกเตอร์แล้วได้เวกเตอร์อีกตัว
ตัวดำเนินการทั้งสองนี้ต้องมีสมบัติต่อไปนี้

1) การบวกเวกเตอร์ต้องมีสมบัติ

1.1) การสลับที่การบวกเวกเตอร์

$$|a\rangle + |b\rangle = |b\rangle + |a\rangle$$

1.2) การเปลี่ยนกลุ่มการบวกเวกเตอร์

$$(|a\rangle + |b\rangle) + |c\rangle = |a\rangle + (|b\rangle + |c\rangle)$$

1.3) การมีอยู่ของเวกเตอร์ศูนย์ $|0\rangle$ ซึ่งมีสมบัติ

$$|a\rangle + |0\rangle = |a\rangle$$

สำหรับเวกเตอร์ $|a\rangle$ ใดๆ

1.4) การมีอยู่ของอินเวอรส์เวกเตอร์ (เวกเตอร์ลบ) สำหรับเวกเตอร์ $|a\rangle$ ใดๆ ซึ่งแทนด้วย $|-a\rangle$ และมีสมบัติคือ

$$|a\rangle + |-a\rangle = |0\rangle$$

2) การคูณสเกล่าร์ต้องมีสมบัติ

2.1) การเปลี่ยนกลุ่มการคูณ

$$\lambda(\sigma|v\rangle) = (\lambda\sigma)|v\rangle$$

2.2) การคูณด้วย 1

3) การบวกเวกเตอร์และการคูณสเกล่าร์มีสมบัติการกระจาย

$$\lambda(|a\rangle + |b\rangle) = \lambda|a\rangle + \lambda|b\rangle$$

$$(\lambda + \sigma)|a\rangle = \lambda|a\rangle + \lambda|a\rangle$$

<u>โดยเวกเตอร์และสเกล่าร์ที่มีตัวดำเนินการการบวกเวกเตอร์และการคูณสเกล่าร์ที่มีสมบัติดังกล่าวนั้นจะเรียกว่าปริภูมิเวกเตอร์ (Vector space)</u>

ตัวอย่าง

จงแสดงว่าจำนวนจริง R เป็นปริภูมิเวกเตอร์บนสเกล่าร์ที่เป็นจำนวนจริง

ตัวอย่าง

จงแสดงว่าจำนวนเชิงซ้อน C เป็นปริภูมิเวกเตอร์บนสเกล่าร์ที่เป็นจำนวนจริง

ตัวอย่าง

จงแสดงว่าจำนวนเชิงซ้อน C เป็นปริภูมิเวกเตอร์บนสเกล่าร์ที่เป็นจำนวนซ้อน

ตัวอย่าง

จงแสดงว่าลูกศรบนกระดาษที่มีกฎการบวกเป็นการวาดหางต่อหัวเป็นปริภูมิเวกเตอร์

ตัวอย่าง

จงแสดงว่าจำนวนจริงสองตัวที่เป็น tuple

$$|x\rangle = (x_1, x_2)$$

ที่มีกฎการบวกเวกเตอร์เป็น

$$|x\rangle + |y\rangle = (x_1, x_2) + (y_1, y_2) \equiv (x_1 + y_1, x_2 + y_2)$$

และกฎการคูณสเกล่าร์เป็น

$$\lambda|x\rangle = \lambda(x_1, x_2) \equiv (\lambda x_1, \lambda x_2)$$

เป็นปริภูมิเวกเตอร์บนสเกล่าร์ที่เป็นจำนวนจริง

ตัวอย่าง

จงแสดงว่าจำนวนจริงจำนวน n ตัวที่เป็น tuple

$$|x\rangle = (x_1, x_2, \dots, x_n)$$

ที่มีกฎการบวกเวกเตอร์เป็น

$$|x\rangle + |y\rangle = (x_1, x_2 \dots, x_n) + (y_1, y_2 \dots, y_n) \equiv (x_1 + y_1, x_2 + y_2 \dots, x_n + y_n)$$

และกฎการคูณสเกล่าร์เป็น

$$\lambda|x\rangle = \lambda(x_1, x_2 \dots, x_n) \equiv (\lambda x_1, \lambda x_2 \dots, \lambda x_n)$$

เป็นปริภูมิเวกเตอร์บนสเกล่าร์ที่เป็นจำนวนจริง

ตัวอย่าง

จงแสดงว่าจำนวนจริงจำนวน ∞ ตัวที่เป็น tuple

$$|x\rangle = (x_1, x_2, \dots)$$

ที่มีกฎการบวกเวกเตอร์เป็น

$$|x\rangle + |y\rangle = (x_1, x_2 \dots) + (y_1, y_2 \dots) \equiv (x_1 + y_1, x_2 + y_2 \dots)$$

และกฎการคูณสเกล่าร์เป็น

$$\lambda|x\rangle = \lambda(x_1, x_2 \dots) \equiv (\lambda x_1, \lambda x_2 \dots)$$

เป็นปริภูมิเวกเตอร์บนสเกล่าร์ที่เป็นจำนวนจริง

ตัวอย่าง

จงแสดงว่าพหุนามดีกรี 2

$$|A\rangle = A(x) = a_0 + a_1 x + a_2 x^2$$

ที่มีกฎการบวกเวกเตอร์เป็นการบวกฟังก์ชันปรกติ

$$|A\rangle + |B\rangle = A(x) + B(x)$$

$$|A\rangle + |B\rangle = a_0 + a_1 x + a_2 x^2 + b_0 + b_1 x + b_2 x^2 = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2$$

และกฎการคูณสเกล่าร์เป็นการคูณฟังก์ชันปรกติ

$$\lambda|A\rangle = \lambda A(x) = \lambda(a_0 + a_1 x + a_2 x^2) \equiv \lambda a_0 + \lambda a_1 x + \lambda a_2 x^2$$

เป็นปริภูมิเวกเตอร์บนสเกล่าร์ที่เป็นจำนวนจริง

ตัวอย่าง

จงแสดงว่าพหุนามดีกรี ∞

$$|A\rangle = A(x) = a_0 + a_1 x + a_2 x^2 + \cdots$$

ที่มีกฎการบวกเวกเตอร์เป็นการบวกฟังก์ชันปรกติ

$$|A\rangle + |B\rangle = A(x) + B(x) = a_0 + a_1 x + a_2 x^2 + \cdots + b_0 + b_1 x + b_2 x^2 + \cdots$$

$$|A\rangle + |B\rangle = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \cdots$$

และกฎการคูณสเกล่าร์เป็นการคูณฟังก์ชันปรกติ

$$\lambda|A\rangle = \lambda A(x) = \lambda(a_0 + a_1 x + a_2 x^2 + \cdots) \equiv \lambda a_0 + \lambda a_1 x + \lambda a_2 x^2 + \cdots$$

เป็นปริภูมิเวกเตอร์บนสเกล่าร์ที่เป็นจำนวนจริง

ตัวอย่าง

จงแสดงว่าฟังก์ชันจำนวนจริง

$$|A\rangle = A(x)$$

ที่มีกฎการบวกเวกเตอร์เป็นการบวกฟังก์ชันปรกติ

$$|A\rangle + |B\rangle = A(x) + B(x)$$

และกฎการคูณสเกล่าร์เป็นการคูณฟังก์ชันปรกติ

$$\lambda|A\rangle = \lambda A(x)$$

เป็นปริภูมิเวกเตอร์บนสเกล่าร์ที่เป็นจำนวนจริง

ตัวอย่าง

จงแสดงว่าฟังก์ชันจำนวนเชิงซ้อน

$$|A\rangle = A(x)$$

ที่มีกฎการบวกเวกเตอร์เป็นการบวกฟังก์ชันปรกติ

$$|A\rangle + |B\rangle = A(x) + B(x)$$

และกฎการคูณสเกล่าร์เป็นการคูณฟังก์ชันปรกติ

$$\lambda|A\rangle = \lambda A(x)$$

เป็นปริภูมิเวกเตอร์บนสเกล่าร์ที่เป็นจำนวนเชิงซ้อน

1.2) ผลรวมเชิงเส้นและความเป็นอิสระเชิงเส้น (Linear Combination and Linearly independence)

ผลรวมเชิงเส้น (linear combination) ของเวกเตอร์คือการคูณสเกล่าร์และบวกเวกเตอร์เข้าด้วยกันเช่น

$$|u\rangle = a|v\rangle + b|w\rangle$$

ในความหมายของผลรวมเชิงเส้นคือเวกเตอร์ $|u\rangle$ สามารถถูกสร้างได้จากเวกเตอร์ $|v\rangle$ และ $|w\rangle$ ดังนั้นเราจะบอกว่า $|u\rangle$ ไม่เป็นอิสระเชิงเส้น (Linearly dependence) กับ $|v\rangle$ และ $|w\rangle$

เวกเตอร์ที่เป็นอิสระเชิงเส้นกันคือเวกเตอร์ที่ไม่สามารถเขียนในรูปของผลรวมเชิงเส้นของกันและกันได้ นั่นคือถ้า

$$c|u\rangle = a|v\rangle + b|w\rangle$$

แล้ว $|u\rangle$ จะเป็นอิสระเชิงเส้นกับ $|v\rangle$ และ $|w\rangle$ ก็ต่อเมื่อ $a = b = c = 0$

ดังนั้นเงื่อนไขของการเป็นอิสระเชิงเส้นคือ

เวกเตอร์ $|u_1\rangle, |u_2\rangle, \dots, |u_n\rangle$ จะเป็นอิสระเชิงเส้น ถ้าความสัมพันธ์

$$\sum_{i=1}^{n} a_i |u_i\rangle = 0$$

มีผลลัพธ์เป็น

$$a_1 = a_2 = \cdots = a_n = 0$$

ตัวอย่าง

จงแสดงว่า

$$|A\rangle = a_0, \qquad |B\rangle = b_1 x, \qquad |C\rangle = c_2 x^2$$

เป็นอิสระเชิงเส้นกัน

ตัวอย่าง

จงแสดงว่า

$$|A\rangle = (1,0,0), \qquad |B\rangle = (0,1,0), \qquad |C\rangle = (0,0,1)$$

เป็นอิสระเชิงเส้นกัน

ตัวอย่าง

จงแสดงว่า

$$|A\rangle = (1,1,1), \qquad |B\rangle = (1,-1,0), \qquad |C\rangle = (1,0,-1)$$

เป็นอิสระเชิงเส้นกัน

1.3) ปริภูมิย่อย, การแผ่ทั่ว และเวกเตอร์ฐาน (Subspace, span, and basis)

ถ้ามีเซตย่อยของปริภูมิเวกเตอร์ที่มีลักษณะปิดภายใต้ผลรวมเชิงเส้นเราจะเรียกเซตย่อยนั้นว่า ปริภูมิย่อย (subspace) นั่นคือ

สำหรับเวกเตอร์สองอันใดๆ $|A\rangle, |B\rangle$ ในปริภูมิย่อยผลรวมเชิงเส้นใดๆ

$$\alpha|A\rangle + \beta|B\rangle$$

จะอยู่ในเซตย่อยนั้นเสมอ

ในอีกทางหนึ่งถ้าเราเลือกเวกเตอร์จากปริภูมิเวกเตอร์ขึ้นมาชุดหนึ่งแล้วทำผลรวมเชิงเส้นทุกทางที่เป็นไปได้เราจะได้เซตแผ่ทั่ว เช่น

สำหรับเวกเตอร์ $|A\rangle, |B\rangle$ ผลรวมเชิงเส้นทุกทางที่เป็นไปได้

$$\alpha|A\rangle + \beta|B\rangle$$

จะเป็นเซตแผ่ทั่ว (span) ของ $|A\rangle, |B\rangle$

เวกเตอร์ฐานคือชุดของเวกเตอร์ที่เป็นอิสระเชิงเส้นและเซตแผ่ทั่วของมันคือปริภูมิเวกเตอร์

ตัวอย่าง

จงหาเซตแผ่ทั่วของ

$$|A\rangle = (1,0), \qquad |B\rangle = (0,1)$$

พร้อมทั้งแสดงว่ามันเป็นปริภูมิย่อย

ตัวอย่าง

จงแสดงว่า

$$|A\rangle = (1,0,0), \qquad |B\rangle = (0,1,0), \qquad |C\rangle = (0,0,1)$$

เป็นเวกเตอร์ฐาน

ตัวอย่าง

จงแสดงว่า

$$|A\rangle = (1,1,1), \qquad |B\rangle = (1,-1,0), \qquad |C\rangle = (1,0,-1)$$

เป็นเวกเตอร์ฐาน

จำนวนเวกเตอร์ฐานที่มีสำหรับปริภูมิเวกเตอร์หนึ่งๆมีค่าคงที่เสมอ และจำนวนเวกเตอร์ฐานคือจำนวนมิติของปริภูมิเวกเตอร์ และสเกล่าร์ที่คูณกับเวกเตอร์ฐานจะถูกเรียกว่าองค์ประกอบของเวกเตอร์ เช่น

$$|v\rangle = \sum_{i=1}^{n} v_i |e_i\rangle$$

เราเรียก $|e_i\rangle$ ว่าเป็นเวกเตอร์ฐานใน n มิติ และ v_i คือองค์ประกอบของเวกเตอร์ $|v\rangle$

ตัวอย่าง

พหุนามดีกรี 2

$$|A\rangle = A(x) = a_0 + a_1 x + a_2 x^2$$

จงตรวจสอบว่า $|e_1\rangle = 1, |e_2\rangle = x, |e_3\rangle = x^2$

ซึ่งเราพบว่ามีจำนวนเวกเตอร์ฐานเท่ากับ 3 *($|e_1\rangle = 1, |e_2\rangle = x, |e_3\rangle = x^2$)* นั่นคือปริภูมิเวกเตอร์นี้มีสามมิติ

ตัวอย่าง

จงแสดงว่าพหุนามดีกรี ∞

$$|A\rangle = A(x) = a_0 + a_1 x + a_2 x^2 + \cdots$$

มีจำนวนมิติเป็นอนันต์

Chapter 1: Vector Space

1.4) ผลคูณภายใน (Inner Product)

ปริภูมิเวกเตอร์ที่เราได้กล่าวถึงในตอนที่แล้วมีโครงสร้างทั่วไปจนอาจจะไม่น่าสนใจและมีประโยชน์สำหรับฟิสิกส์สักเท่าไหร่ สิ่งที่เรามักสนใจเกี่ยวกับเวกเตอร์ในฟิสิกส์คือขนาดและทิศทางซึ่งต้องการโครงสร้างทางคณิตศาสตร์ที่เรียกว่าผลคูณภายใน (Inner Product) โดยเรารู้จักผลคูณนี้ในชื่อผลคูณดอท (Dot Product) ที่เรียนไปแล้วในเทอมที่แล้ว

โดยโครงสร้างของการทำผลคูณดอทคือการคูณเวกเตอร์สองตัวเพื่อได้ตัวเลขที่เป็นสเกล่าร์ ดังนั้นเราจะเขียนด้วยสัญลักษณ์

$$g(|a\rangle, |b\rangle) \in \mathbb{R}$$

สมบัติของผลคูณนี้ต้องมีลักษณะเชิงเส้น นั่นคือ

$$g(\alpha|a\rangle + \beta|b\rangle, |c\rangle) = \alpha g(|a\rangle, |c\rangle) + \beta g(|b\rangle, |c\rangle)$$

และต้องมีนิยามของความยาวซึ่งมีค่ามากกว่าศูนย์ นั่นคือ

$$g(|a\rangle, |a\rangle) = |a|^2 \geq 0$$

และต้องมีเพียงเวกเตอร์เดียว (เวกเตอร์ศูนย์) ที่มีความยาวเป็นศูนย์

$$g(|a\rangle, |a\rangle) = 0 \iff |a\rangle = |0\rangle$$

อย่างไรก็ตามหากเราต้องการจะขยายโครงสร้างนี้ให้รองรับกับปริภูมิเวกเตอร์เชิงซ้อนด้วย (Complex vector spaces) เราจะพบว่าปัญหาของการขยายผลนี้คือถ้าเราใช้สมบัติเชิงเส้นเราจะพบว่า

$$g(i|a\rangle, i|a\rangle) = i^2 g(|a\rangle, |a\rangle) = -g(|a\rangle, |a\rangle)$$

ซึ่งจะทำให้ความยาวนี้เป็นลบ (เพราะฝั่งใดฝั่งหนึ่งต้องเป็นค่าลบ) เพื่อที่จะแก้ปัญหานี้เราจะให้ความเป็นเชิงเส้นสำหรับจำนวนเชิงซ้อนมีลักษณะเป็น complex conjugate สำหรับเวกเตอร์ตัวแรกนั่นคือ

$$g(\alpha|a\rangle + \beta|b\rangle, |c\rangle) = \alpha^* g(|a\rangle, |c\rangle) + \beta^* g(|b\rangle, |c\rangle)$$

และเพื่อให้สอดคล้องกันเราต้องให้สมบัติความสมมาตรของผลคูณภายในมี complex conjugate มาเกี่ยวข้องด้วย

$$g(|a\rangle, |b\rangle) = g(|b\rangle, |a\rangle)^*$$

ต่อจากนี้ไปเราจะใช้สัญลักษณ์การทำผลคูณภายในของดิแรก (Dirac bracket notation) โดยเวกเตอร์ที่เราใช้จะถูกเรียกว่าเคทเวกเตอร์ (ket vector, $|b\rangle$) และเวกเตอร์ที่ใช้ในการทำผลคูณภายในกับเคทเวกเตอร์จะถูกเรียกว่าบราเวกเตอร์ (bra vector,$\langle a|$) โดยผลคูณภายในเกิดจากการประกบกันของบราและเคทเวกเตอร์ (bra + ket = bra(c)ket)

$$g(|a\rangle, |b\rangle) \equiv \langle a|b\rangle \in \mathbb{C}$$

นิยาม ผลคูณภายใน (Inner Product) ของเวกเตอร์ และ ในปริภูมิเวกเตอร์ คือจำนวนเชิงซ้อนซึ่งมีสมบัติต่อไปนี้

1. $\langle a|b\rangle = \langle b|a\rangle^*$
2. $\langle a|(\beta|b\rangle + \gamma|c\rangle) = \beta\langle a|b\rangle + \gamma\langle a|c\rangle$
3. $\langle a|a\rangle \geq 0$ และ $\langle a|a\rangle = 0$ ก็ต่อเมื่อ $|a\rangle = |0\rangle$

โดยสมบัติสุดท้ายนี้เรียกว่าสมบัติ positive definite ซึ่งในฟิสิกส์บางสาขา (เช่นสัมพัทธภาพพิเศษ) อาจจะไม่มีความจำเป็นต้องใช้สมบัตินี้ โดยเราจะเรียกผลคูณภายในที่ไม่มีสมบัติ positive definite ว่า pseudo-Riemannian inner product

ตัวอย่าง จงแสดงว่าในปริภูมิเวกเตอร์ complex 2-tuple ที่มีการคูณ
$|a\rangle = (a_1, a_2)$ และ $|b\rangle = (b_1, b_2)$

$$\langle a|b\rangle = a_1^* b_1 + a_2^* b_2$$

เป็นผลคูณภายใน

ตัวอย่าง จงแสดงว่าในปริภูมิเวกเตอร์ Real 3-tuple ที่มีการคูณ

$|a\rangle = (a_1, a_2, a_3)$ และ $|b\rangle = (b_1, b_2, b_3)$

$$\langle a|b\rangle = a_1 b_1 + a_2 b_2 + a_3 b_3$$

เป็นผลคูณภายใน

(สังเกตว่าผลคูณภายในนี้คือการทำผลคูณดอทในเวกเตอร์สามมิติแบบที่เราคุ้นเคยกัน)

ตัวอย่าง สำหรับปริภูมิเวกเตอร์เชิงซ้อนของพหุนามดีกรีอนันต์

$$|a\rangle = a_0 + a_1 x + a_2 x^2 + \cdots = a(x)$$

โดยที่ $a_0, a_1, \ldots \in \mathbb{C}$

จงแสดงว่าผลคูณนี้

$$\langle a|b\rangle \equiv \int_{x_i}^{x_f} a^*(x)\, b(x) dx$$

เป็นผลคูณภายใน

1.5 การตั้งฉาก (Orthogonality)

เวกเตอร์ $|a\rangle$ และ $|b\rangle$ เป็นเวกเตอร์ที่ตั้งฉากกัน (orthogonal) ก็ต่อเมื่อ $\langle a|b\rangle = 0$

เวกเตอร์ เป็นเวกเตอร์ที่มีขนาดเท่ากับ 1 (normal) ก็ต่อเมื่อ $\langle a|a\rangle = 1$

เวกเตอร์ฐานที่มีสมบัติ orthonormal คือเวกเตอร์ฐาน $\{|e_i\rangle\}$ เมื่อ $i = 1, \dots, n$ ที่สอดคล้องกับ

$$\langle e_i|e_j\rangle = \delta_{ij} \equiv \begin{cases} 1 & \text{ถ้า } i = j \\ 0 & \text{ถ้า } i \neq j \end{cases}$$

โดยที่ δ_{ij} คือสัญลักษณ์โครเนกเกอร์เดลต้า (Kronecker delta)

ตัวอย่าง จงพิจารณาว่าเวกเตอร์ตัวใดตั้งฉากกันบ้างภายใต้ผลคูณภายใน

$$\langle a|b\rangle = a_1 b_1 + a_2 b_2 + a_3 b_3$$

1. $|a_1\rangle = (1,0,-1)$
2. $|a_2\rangle = (1,0,1)$
3. $|a_3\rangle = (-1,0,1)$
4. $|a_4\rangle = (1,-2,-1)$
5. $|a_5\rangle = (0,1,0)$

ตัวอย่าง จงพิจารณาว่าเวกเตอร์ตัวใดตั้งฉากกันบ้างภายใต้ผลคูณภายใน

$$\langle f|g\rangle = \frac{2}{L}\int_0^L f(x)g(x)dx$$

1. $|f_1\rangle = \sin\left(\frac{2\pi x}{L}\right)$
2. $|f_2\rangle = \cos\left(\frac{2\pi x}{L}\right)$
3. $|f_3\rangle = \sin\left(\frac{4\pi x}{L}\right)$
4. $|f_4\rangle = \cos\left(\frac{4\pi x}{L}\right)$

ตัวอย่าง จงหาขนาดของเวกเตอร์ต่อไปนี้ภายใต้ผลคูณภายใน

$$\langle f|g\rangle = \frac{2}{L}\int_0^L f(x)g(x)dx$$

1. $|f_1\rangle = \sin\left(\frac{2\pi x}{L}\right)$
2. $|f_2\rangle = \cos\left(\frac{2\pi x}{L}\right)$
3. $|f_3\rangle = \sin\left(\frac{4\pi x}{L}\right)$
4. $|f_4\rangle = \cos\left(\frac{4\pi x}{L}\right)$

1.6 กระบวนการกรัม-ชมิดท์ (Gram-Schmidt process)

ในหลายๆ ครั้งเราต้องเริ่มทำงานกับเวกเตอร์ฐานที่ไม่เป็น orthonormal ซึ่งอาจจะไม่สะดวกกับการคำนวณบางประเภท เพื่อเป็นการทำให้การคำนวณง่ายขึ้น เราอาจจะต้องหากระบวนการที่ทำให้เวกเตอร์ฐานใดๆเปลี่ยนกลับมาเป็นเวกเตอร์ฐานที่เป็น orthonormal กระบวนการนี้เราเรียกว่า Gram-Schmidt process

เราจะเริ่มจากกรณีเวกเตอร์ฐานสามตัวก่อน

$$\{|a_1\rangle, |a_2\rangle, |a_3\rangle\}$$

ขั้นตอนที่ 1: Normalize เวกเตอร์ตัวแรก

เราจะยึดเวกเตอร์ตัวแรกเป็นหลัก ดังนั้นเราจะทำให้ขนาดมีค่าเท่ากับ 1 ก่อนโดยการหารด้วย

$$|e_1\rangle = \frac{|a_1\rangle}{\sqrt{\langle a_1|a_1\rangle}}$$

ซึ่งนี่จะทำให้ $\langle e_1|e_1\rangle = \frac{\langle a_1|a_1\rangle}{\langle a_1|a_1\rangle} = 1$

ขั้นตอนที่ 2: หักองค์ประกอบของเวกเตอร์ตัวที่สองในทิศทางของเวกเตอร์ตัวแรก

ส่วนของเวกเตอร์ตัวที่สองที่อยู่ในทิศทางของเวกเตอร์ตัวแรกคือ

$$\langle e_1|a_2\rangle \cdot |e_1\rangle$$

ซึ่งถ้าเราต้องการให้เวกเตอร์ที่สร้างใหม่ตั้งฉากกับเวกเตอร์ตัวแรก เราต้องทำการลบองค์ประกอบในแกนของเวกเตอร์ตัวแรกออกนั่นคือ

$$|e_2'\rangle = |a_2\rangle - \langle e_1|a_2\rangle \cdot |e_1\rangle$$

ซึ่งจะทำให้เราได้เวกเตอร์ที่ตั้งฉากกับเวกเตอร์ตัวแรก แต่ยังมีขนาดไม่เท่ากับ 1

ขั้นตอนที่ 3: Normalize เวกเตอร์ตัวที่สอง

เพื่อทำให้มันเป็น orthonormal เราต้องทำการ normalize

$$|e_2\rangle = \frac{|e_2{}'\rangle}{\sqrt{\langle e_2'|e_2'\rangle}}$$

ขั้นตอนที่ 4: หักองค์ประกอบของเวกเตอร์ตัวที่สามในทิศทางของเวกเตอร์สองตัวแรก

ส่วนของเวกเตอร์ตัวที่สามที่อยู่ในทิศทางของเวกเตอร์สองตัวแรกคือ

$$\langle e_1|a_3\rangle \cdot |e_1\rangle + \langle e_2|a_3\rangle \cdot |e_2\rangle$$

ซึ่งถ้าเราต้องการให้เวกเตอร์ที่สร้างใหม่ตั้งฉากกับเวกเตอร์สองตัวแรก เราต้องทำการลบองค์ประกอบในแกนของเวกเตอร์สองตัวแรกออกนั่นคือ

$$|e_3'\rangle = |a_3\rangle - \langle e_1|a_3\rangle \cdot |e_1\rangle - \langle e_2|a_3\rangle \cdot |e_2\rangle$$

ขั้นตอนที่ 5: Normalize เวกเตอร์ตัวที่สาม

$$|e_3\rangle = \frac{|e_3{}'\rangle}{\sqrt{\langle e_3'|e_3'\rangle}}$$

ผลลัพธ์คือเวกเตอร์ ซึ่งเป็น orthonormal กันทั้งหมด ซึ่งในกรณีที่มีจำนวนเวกเตอร์มากกว่านี้ เราก็สามารถทำกระบวนการนี้ซ้ำๆเพื่อหาเวกเตอร์ที่เป็น orthonormal กันทั้งหมดได้

ตัวอย่าง จงแสดงว่า $|e_2\rangle$ และ $|e_3'\rangle$ ตั้งฉากกัน

ตัวอย่าง จงหาเวกเตอร์ฐานที่ orthonormal จากเวกเตอร์ฐานต่อไปนี้

$|a_1\rangle = (0,1,0), |a_2\rangle = (1,1,1), |a_3\rangle = (1,0,-1)$

1.7 อสมการของชวาร์ส์ (Schwarz inequality)

สำหรับเวกเตอร์ $|a\rangle$ และ $|b\rangle$ ใดๆ อสมการต่อไปนี้เป็นจริงเสมอ

$$\langle a|a\rangle\langle b|b\rangle \geq |\langle a|b\rangle|^2$$

ซึ่งอสมการนี้จะกลายเป็นสมการเมื่อ และ อยู่ในทิศทางเดียวกันนั่น คือ

$$|a\rangle = \lambda|b\rangle \rightarrow \langle a|a\rangle\langle b|b\rangle = |\langle a|b\rangle|^2$$

ตัวอย่าง จงพิสูจน์อสมการนี้โดยการพิจารณาเวกเตอร์

$$|c\rangle = |b\rangle - \left(\frac{\langle a|b\rangle}{\langle a|a\rangle}\right)|a\rangle$$

และใช้เงื่อนไขที่ว่า $\langle c|c\rangle \geq 0$

แบบฝึกหัด 1

1) Matrix vector space 2x2

1.1) จงแสดงว่าแมทริกซ์ 2x2 ที่เป็นจำนวนจริง

$$|v\rangle = \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix}$$

ที่มีการบวกเวกเตอร์เป็น

$$|v\rangle + |w\rangle = \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix} + \begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix} \equiv \begin{pmatrix} v_{11} + w_{11} & v_{12} + w_{12} \\ v_{21} + w_{21} & v_{22} + w_{22} \end{pmatrix}$$

และการคูณสเกล่าร์

$$\lambda|v\rangle = \begin{pmatrix} \lambda v_{11} & \lambda v_{12} \\ \lambda v_{21} & \lambda v_{22} \end{pmatrix}$$

เป็นปริภูมิเวกเตอร์จริง (real vector space)

1.2) จงแสดงว่าเวกเตอร์ต่อไปนี้เป็นอิสระเชิงเส้นกัน

$$|v_1\rangle = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, |v_2\rangle = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, |v_1\rangle = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, |v_1\rangle = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

1.3) จงแสดงว่าเวกเตอร์ต่อไปนี้เป็นอิสระเชิงเส้นกัน

$$|v_1\rangle = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, |v_2\rangle = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, |v_1\rangle = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, |v_1\rangle = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

1.4) ถ้ากำหนดให้ product เป็น

$$\langle v|w\rangle = Tr\left[\begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix}^T \begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix}\right] = v_{11}w_{11} + v_{12}w_{12} + v_{21}w_{21} + v_{22}w_{22}$$

จงแสดงว่า product นี้มีสมบัติเป็น inner product

1.5) จาก 1.4) จงแสดงว่าเวกเตอร์ใน 1.2) เป็น orthonormal basis

1.6) จาก 1.4) จงใช้ inner product นี้ทำให้เวกเตอร์ฐานใน 1.3) เป็น orthonormal ผ่านทาง

Gram-Schmidt process

2) ผลรวมเชิงเส้นและอิสระเชิงเส้น

2.1) จงพิจารณาว่าเวกเตอร์ใดต่อไปนี้เป็นผลรวมเชิงเส้นของ

$$|f\rangle = f(x) = -2 + x + 6x^2, \qquad |g\rangle = g(x) = 3 - 4x + x^2$$

- $h_1(x) = 1 - 3x + 7x^2$
- $h_2(x) = -1 - 2x + 12x^2$
- $h_3(x) = 3 - 6x + 2x^2 + 6x^3$
- $h_4(x) = -8 + 9x + 4x^2$

2.2) จงพิจารณาว่าเวกเตอร์ใดต่อไปนี้เป็นอิสระเชิงเส้นกับ

$$|f\rangle = (1,2,0), \qquad |g\rangle = (3,2,1)$$

- $|h_1\rangle = (1,0,0)$
- $|h_2\rangle = (1,2,1)$
- $|h_3\rangle = (4,4,1)$
- $|h_4\rangle = (7,6,2)$

3. ผลคูณภายในและการตั้งฉาก

3.1) จงหาผลคูณภายใน

$$|a\rangle = (a_1, a_2, a_3), \qquad |b\rangle = (b_1, b_2, b_3)$$

$$\langle a|b\rangle = a_1^* b_1 + a_2^* b_2 + a_3^* b_3$$

ของเวกเตอร์ต่อไปนี้

- $|a\rangle = (i, 1+i, 1-i)$, $|b\rangle = (1,1,-1)$
- $|a\rangle = (0, i, -i)$, $|b\rangle = (i, i, 1)$

3.2) จงหาคู่ของเวกเตอร์ต่อไปนี้ที่ตั้งฉากกัน

- $|a_1\rangle = (i, 0, -i)$

- $|a_2\rangle = (i, 1, i)$

- $|a_3\rangle = (0,1, -i)$

- $|a_4\rangle = (1,1, -i)$

3.3) จงหาผลคูณภายใน

$$\langle f|g\rangle = \frac{2}{L}\int_0^L f^*(x)g(x)dx$$

ของเวกเตอร์ต่อไปนี้

- $f_1(x) = e^{\frac{2\pi xi}{L}}, \; f_2(x) = e^{\frac{3\pi xi}{L}}$

- $f_1(x) = e^{\frac{2\pi xi}{L}}, \; f_2(x) = e^{-\frac{\pi xi}{L}}$

- $f_1(x) = e^{\frac{\pi xi}{L}}, \; f_2(x) = e^{-\frac{\pi xi}{L}}$

3.4) จงหาเวกเตอร์ต่อไปนี้ที่ตั้งฉากกัน

- $f(x) = e^{\frac{2\pi xi}{L}}$

- $f(x) = e^{\frac{\pi xi}{L}}$

- $f(x) = e^{\frac{3\pi xi}{L}}$

- $f(x) = e^{\frac{4\pi xi}{L}}$

Chapter 2: Operators

2.1) การแปลงเชิงเส้นและตัวดำเนินการเชิงเส้น (Linear Transformation and Linear Operator)

การส่ง (mapping) คือความเชื่อมโยงจากสมาชิกในเซตหนึ่งไปยังสมาชิกในอีกเซตหนึ่ง โดยความเชื่อมโยงนั้นต้องเชื่อมไปหาสมาชิกเพียงตัวเดียวในเซตที่สองเท่านั้น ตัวอย่างเช่น

1. $f: \mathbb{R} \to \mathbb{R}$ โดยที่ $f(x) = x^2$
2. $g: \mathbb{R}^2 \to \mathbb{R}$ โดยที่ $g(x, y) = x^2 + y^2$
3. $h: \mathbb{R} \to \mathbb{R}^2$ โดยที่ $h(x) = (x + 1, x^2)$

การแปลงเชิงเส้น (Linear transformation) คือการส่งจากปริภูมิเวกเตอร์หนึ่งไปยังอีกปริภูมิเวกเตอร์หนึ่ง

$$\hat{T}: \mathbb{V} \to \mathbb{W}$$

ซึ่งมีสมบัติเชิงเส้น

$$\hat{T}(\alpha|a\rangle + \beta|b\rangle) = \alpha\hat{T}(|a\rangle) + \beta\hat{T}(|b\rangle)$$

และถ้าการแปลงเชิงเส้นนั้นเป็นการส่งไปยังปริภูมิเวกเตอร์เดิม $\hat{T}: \mathbb{V} \to \mathbb{V}$ เราจะเรียกว่าตัวดำเนินการเชิงเส้น (linear operator)

เพื่อความสะดวกเรามักจะไม่ใส่วงเล็บในตัวกระทำการนั่นคือ

$$\hat{T}(|a\rangle) \equiv \hat{T}|a\rangle$$

ตัวอย่าง จงแสดงว่าการสลับที่ของ 2-tuple เป็นตัวดำเนินการเชิงเส้น

$$|a\rangle = (a_1, a_2) \to \hat{T}|a\rangle = (a_2, a_1)$$

ตัวอย่าง จงแสดงว่าการกลับด้านของ 2-tuple เป็นตัวดำเนินการเชิงเส้น

$$|a\rangle = (a_1, a_2) \to \hat{T}|a\rangle = (-a_1, -a_2)$$

ตัวอย่าง จงแสดงว่าการทำโปรเจคชัน (projection) บนแกนของ 2-tuple เป็นตัวดำเนินการเชิงเส้น

$$|a\rangle = (a_1, a_2) \to \hat{T}|a\rangle = (a_1, 0)$$

ตัวอย่าง จงแสดงว่าการทำอนุพันธ์เป็นตัวดำเนินการเชิงเส้น

$$|a\rangle = (a_0, a_1, a_2, \dots, a_n) \to \hat{T}|a\rangle = (a_1, 2a_2, 3a_3, \dots, na_n, 0)$$

(หรืออาจมองเป็น)

$$|a\rangle = a(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

$$\to \hat{T}|a\rangle = \frac{da(x)}{dx} = a_1 + 2a_2 x + 3a_3 x^2 + \cdots + na_n x^{n-1}$$

2.2) แมทริกซ์ตัวแทน (Matrix representation)

เช่นเดียวกับการเขียนเวกเตอร์ด้วย column vector (tuple) เราสามารถเขียนตัวดำเนินการเชิงเส้น (linear operator) ด้วย matrix ได้

พิจารณาเวกเตอร์หนึ่งซึ่งอยู่ในปริภูมิเวกเตอร์ 2 มิติที่มีเวกเตอร์ฐานเป็น $\{|e_1\rangle, |e_2\rangle\}$

$$|v\rangle = v_1|e_1\rangle + v_2|e_2\rangle$$

ถ้าให้ตัวดำเนินการเชิงเส้นกระทำกับเวกเตอร์ตัวนี้เราจะได้

$$\hat{T}|v\rangle = \hat{T}(v_1|e_1\rangle + v_2|e_2\rangle) = v_1\hat{T}|e_1\rangle + v_2\hat{T}|e_2\rangle$$

ซึ่งถ้ากำหนดให้ตัวดำเนินการเชิงเส้นกระทำบนเวกเตอร์ฐานทั้งสองแล้วได้เป็นเวกเตอร์ต่อไปนี้

$$\hat{T}|e_1\rangle = T_{11}|e_1\rangle + T_{21}|e_2\rangle$$

$$\hat{T}|e_2\rangle = T_{12}|e_1\rangle + T_{22}|e_2\rangle$$

เราจะได้ว่า

$$\hat{T}|v\rangle = v_1(T_{11}|e_1\rangle + T_{21}|e_2\rangle) + v_2(T_{12}|e_1\rangle + T_{22}|e_2\rangle)$$

$$\hat{T}|v\rangle \equiv |v'\rangle = v_1'|e_1\rangle + v_2'|e_2\rangle = (T_{11}v_1 + T_{12}v_2)|e_1\rangle + (T_{21}v_1 + T_{22}v_2)|e_2\rangle$$

ซึ่งเมื่อเขียนเวกเตอร์ใหม่ในรูปขององค์ประกอบจะได้ว่า

$$v_1' = (T_{11}v_1 + T_{12}v_2)$$

$$v_2' = (T_{21}v_1 + T_{22}v_2)$$

แทนในรูปแมทริกซ์

$$\boxed{|v'\rangle = \hat{T}|v\rangle \to \tilde{v}' = \tilde{T}\tilde{v} \to \begin{pmatrix} v_1' \\ v_2' \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}}$$

นั่นคือเราสามารถแทนเวกเตอร์ด้วยองค์ประกอบของเวกเตอร์ดังนี้

$$|v\rangle = v_1|e_1\rangle + v_2|e_2\rangle \to \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

$$|v'\rangle = v_1'|e_1\rangle + v_2'|e_2\rangle \to \begin{pmatrix} v_1' \\ v_2' \end{pmatrix}$$

และตัวดำเนินการเชิงเส้นถูกแทนด้วยแมทริกซ์จัตุรัส (square matrix) ดังนี้

$$\hat{T} \to \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix}$$

ในกรณีที่เป็นมิติใดๆเรามักจะเขียนแทนด้วยเครื่องหมาย summation

$$|v\rangle = v_1|e_1\rangle + v_2|e_2\rangle + \cdots v_n|e_n\rangle = \sum_{i=1}^{n} v_i|e_i\rangle$$

และ

$$\hat{T}|e_i\rangle = \sum_{j=1}^{n} T_{ji}|e_j\rangle$$

$$\boxed{v_i = \sum_{j=1}^{n} T_{ij}v_j}$$

หลังจากนี้เป็นต้นไป เราจะพบการ summation บ่อย และทุกครั้งที่เราพบมันจะมาพร้อมกับดัชนี (index) ที่ซ้ำกันและเป็น dummy index เสมอ ดังนั้นเพื่อเป็นการประหยัดเวลา เราจะใช้ข้อตกลงผลรวมของไอน์สไตน์ Einstein summation convention นั่นคือ

เมื่อเจอดัชนีที่ซ้ำกันเราจะถือว่ามีการใช้ผลรวมโดยไม่ต้องเขียนสัญลักษณ์ผลรวม

เช่น

$$v_i = \sum_{j=1}^{n} T_{ij}v_j \rightarrow v_i = T_{ij}v_j$$

ตัวอย่าง จงแสดงว่าตัวแทนของตัวดำเนินการเชิงเส้นที่เป็นการหมุนไปด้วยมุม θ คือ

$$\hat{T} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

ตัวอย่าง จงแสดงว่าตัวแทนของตัวดำเนินการเชิงเส้นที่เป็นการสะท้อนรอบแกน $x = y$ คือ

$$\hat{T} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

ตัวอย่าง จงแสดงว่าตัวแทนของตัวดำเนินการเชิงเส้นที่เป็นการสะท้อนรอบแกน y คือ

$$\hat{T} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

การทำผลคูณภายในก็สามารถเขียนในรูปของแมทริกซ์ได้เช่นกัน ตัวอย่างเช่น

$$\langle a|b\rangle = a_i b_i \rightarrow \tilde{a}^T\tilde{b} = \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix}\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

ซึ่งเราอาจจะแทรกตรงกลางด้วยแมทริกซ์เอกลักษณ์ (Identity matrix)

$$\tilde{a}^T\tilde{b} = \tilde{a}^T\mathbb{I}\tilde{b} = \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix}\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

ดังนั้นในกรณีทั่วไปเราอาจจะเขียนผลคูณภายในได้เป็น

$$\langle a|b\rangle = \tilde{a}^T\tilde{G}\tilde{b} = \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix}\begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{pmatrix}\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

โดยที่เราเรียกแมทริกซ์ว่าเมทริกแมทริกซ์ (metric matrix)

$$\tilde{G} = \begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{pmatrix}$$

ตัวอย่าง การหาผลคูณภายในของเวกเตอร์สี่มิติในทฤษฎีสัมพัทธภาพพิเศษสามารถเขียนได้เป็น

$$\langle A|B\rangle = t_A t_B - x_A x_B - y_A y_B - z_A z_B$$

ซึ่งผลคูณภายในนี้เป็นแบบ pseudo-Riemannian จงหาเมทริกแมทริกซ์ของผลคูณภายในนี้

Chapter 2: Operators

2.3) การดำเนินการบนแมทริกซ์ (Matrix Operations)

สิ่งที่น่าสนใจเกี่ยวตัวดำเนินการเชิงเส้นมันสามารถใช้เพื่อกระทำกับเวกเตอร์หลายๆครั้งได้นั่นคือ ถ้าเราต้องการดำเนินการด้วยตัวดำเนินการ $\hat{T}$ และ $\hat{S}$ เราจะได้ว่า

$$\hat{T}\big(\hat{S}|v\rangle\big) = \hat{T}\hat{S}|v\rangle$$

ตัวดำเนินการทั้งสองตัวต่อเนื่องกันนี้คิดเป็นตัวดำเนินการตัวเดียวกันได้

$$\hat{T}\hat{S}|v\rangle = \hat{U}|v\rangle \rightarrow \hat{T}\hat{S} = \hat{U}$$

ซึ่งในรูปของแมทริกซ์ตัวแทนเราจะได้เป็นการคูณของแมทริกซ์

$$\begin{pmatrix} T_{11} & \cdots & T_{1n} \\ \vdots & \ddots & \vdots \\ T_{n1} & \cdots & T_{nn} \end{pmatrix}\begin{pmatrix} S_{11} & \cdots & S_{1n} \\ \vdots & \ddots & \vdots \\ S_{n1} & \cdots & S_{nn} \end{pmatrix} = \begin{pmatrix} U_{11} & \cdots & U_{1n} \\ \vdots & \ddots & \vdots \\ U_{n1} & \cdots & U_{nn} \end{pmatrix}$$

ในรูปของแมทริกซ์จัตุรัสขนาด 2x2 เราจะได้

$$\begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix}\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} = \begin{pmatrix} T_{11}S_{11} + T_{12}S_{21} & T_{11}S_{12} + T_{12}T_{22} \\ T_{21}S_{11} + T_{22}T_{21} & T_{21}S_{12} + T_{22}T_{22} \end{pmatrix} = \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix}$$

และในรูปของแมทริกซ์ขนาดใดๆเราจะเขียนในรูปขององค์ประกอบ

$$\boxed{U_{ij} = T_{ik}S_{kj}}$$

ตัวอย่าง จงพิจารณาแมทริกซ์การหมุนด้วยมุม θ และมุม ϕ ต่อเนื่องกัน เพื่อแสดงว่า

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} = \begin{pmatrix} \cos(\theta+\phi) & -\sin(\theta+\phi) \\ \sin(\theta+\phi) & \cos(\theta+\phi) \end{pmatrix}$$

ข้อสังเกตคือการคูณแมทริกซ์โดยทั่วไปไม่มีสมบัติในการสลับที่กันนั่นคือ

$$\hat{T}\hat{S} \neq \hat{S}\hat{T}$$

$$\begin{pmatrix} T_{11} & \cdots & T_{1n} \\ \vdots & \ddots & \vdots \\ T_{n1} & \cdots & T_{nn} \end{pmatrix}\begin{pmatrix} S_{11} & \cdots & S_{1n} \\ \vdots & \ddots & \vdots \\ S_{n1} & \cdots & S_{nn} \end{pmatrix} \neq \begin{pmatrix} S_{11} & \cdots & S_{1n} \\ \vdots & \ddots & \vdots \\ S_{n1} & \cdots & S_{nn} \end{pmatrix}\begin{pmatrix} T_{11} & \cdots & T_{1n} \\ \vdots & \ddots & \vdots \\ T_{n1} & \cdots & T_{nn} \end{pmatrix}$$

แมทริกซ์เอกลักษณ์ (Identity matrix) คือแมทริกซ์ที่มีแนวทแยงมีค่าเป็น 1 และตำแหน่งอื่นมีค่าเป็น 0 ทั้งหมด

$$\mathbb{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

ซึ่งสังเกตได้ว่าองค์ประกอบของแมทริกซ์เอกลักษณ์นี้คือสัญลักษณ์ Kronecker delta

$$\boxed{\mathbb{I}_{ij} = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}}$$

ตัวอย่าง จงแสดงว่าเพาลีแมทริกซ์ (Pauli matrices) ไม่มีสมบัติการสลับที่การคูณ

$$\hat{\sigma}_1 \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \hat{\sigma}_2 \equiv \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \hat{\sigma}_3 \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

การทำ transpose คือการสลับตำแหน่งขององค์ประกอบในแมทริกซ์โดยเป็นการสลับแถวและหลักกันเช่น

$$\begin{pmatrix} S_{11} & \cdots & S_{1n} \\ \vdots & \ddots & \vdots \\ S_{n1} & \cdots & S_{nn} \end{pmatrix}^T = \begin{pmatrix} S_{11} & \cdots & S_{n1} \\ \vdots & \ddots & \vdots \\ S_{1n} & \cdots & S_{nn} \end{pmatrix}$$

$$S_{ij}^T = S_{ji}$$

แมทริกซ์สมมาตร (symmetric) คือแมทริกซ์ที่ไม่เปลี่ยนแปลงภายใต้การทำ transpose นั่นคือ

$$\boxed{\hat{S}^T = \hat{S}}$$

สังเกตว่า

$$\left(\hat{A}^T\right)^T_{ij} = \hat{A}^T_{ji} = \hat{A}_{ij} \rightarrow \left(\hat{A}^T\right)^T = A$$

$$\left(\hat{A}+\hat{B}\right)^T_{ij} = \left(\hat{A}+\hat{B}\right)_{ji} = \hat{A}_{ji} + \hat{B}_{ji} = \hat{A}^T_{ij} + \hat{B}^T_{ij} \rightarrow \left(\hat{A}+\hat{B}\right)^T = \hat{A}^T + \hat{B}^T$$

$$\left(\hat{A}\hat{B}\right)^T_{ij} = \left(\hat{A}\hat{B}\right)_{ji} = \hat{A}_{jk}\hat{B}_{ki} = \hat{A}^T_{kj}\hat{B}^T_{ik} = \hat{B}^T_{ik}\hat{A}^T_{kj} = \left(\hat{B}^T\hat{A}^T\right)_{ij} \rightarrow \left(\hat{A}\hat{B}\right)^T = \hat{B}^T\hat{A}^T$$

ในกรณีที่องค์ประกอบของแมทริกซ์เป็นจำนวนเชิงซ้อน การทำ adjoint หรือ dagger คือการทำ transpose และทำ complex conjugate พร้อมกัน นั่นคือ

$$\begin{pmatrix} S_{11} & \cdots & S_{1n} \\ \vdots & \ddots & \vdots \\ S_{n1} & \cdots & S_{nn} \end{pmatrix}^\dagger = \begin{pmatrix} S_{11} & \cdots & S_{1n} \\ \vdots & \ddots & \vdots \\ S_{n1} & \cdots & S_{nn} \end{pmatrix}^{T^*} = \begin{pmatrix} S^*_{11} & \cdots & S^*_{1n} \\ \vdots & \ddots & \vdots \\ S^*_{n1} & \cdots & S^*_{nn} \end{pmatrix}^T = \begin{pmatrix} S^*_{11} & \cdots & S^*_{n1} \\ \vdots & \ddots & \vdots \\ S^*_{1n} & \cdots & S^*_{nn} \end{pmatrix}$$

$$S^\dagger_{ij} = S^*_{ji}$$

แมทริกซ์ที่ไม่เปลี่ยนแปลงภายใต้การทำ adjoint จะถูกเรียกว่าเฮอมิเชียนแมทริกซ์ (Hermitian matrix)

$$\boxed{\hat{H}^\dagger = \hat{H}}$$

สังเกตว่า

$$\left(\hat{A}^\dagger\right)^\dagger = A, \qquad \left(\hat{A}+\hat{B}\right)^\dagger = \hat{A}^\dagger + \hat{B}^\dagger, \qquad \left(\hat{A}\hat{B}\right)^\dagger = \hat{B}^\dagger\hat{A}^\dagger$$

ตัวอย่าง จงแสดงว่าเพาลีแมทริกซ์เป็นเฮอมิเชียนแมทริกซ์

$$\hat{\sigma}^\dagger{}_1 = \hat{\sigma}_1, \qquad \hat{\sigma}^\dagger{}_2 = \hat{\sigma}_2, \qquad \hat{\sigma}^\dagger{}_3 = \hat{\sigma}_3$$

แมทริกซ์อินเวอร์ส (Inverse matrix) คือแมทริกซ์ส่วนกลับที่คูณกับแมทริกซ์เดิมแล้วได้แมทริกซ์เอกลักษณ์

$$\hat{A}\hat{A}^{-1} = \mathbb{I}$$

ตัวอย่าง จงแสดงว่าเพาลีแมทริกซ์เป็นอินเวอร์สของตัวเอง (involutory matrix) นั่นคือ

$$\hat{\sigma}_1\hat{\sigma}_1 = \mathbb{I}, \qquad \hat{\sigma}_2\hat{\sigma}_2 = \mathbb{I}, \qquad \hat{\sigma}_3\hat{\sigma}_3 = \mathbb{I}$$
$$\hat{\sigma}_1^{-1} = \hat{\sigma}_1, \qquad \hat{\sigma}_2^{-1} = \hat{\sigma}_2, \qquad \hat{\sigma}_3^{-1} = \hat{\sigma}_3$$

2.4 ดีเทอมิแนนท์และจาโคเบียน (Determinant and Jacobian)

พิจารณาการแก้สมการ

$$a_{11}x + a_{12}y = b_1$$
$$a_{21}x + a_{22}y = b_2$$

เพื่อที่จะกำจัดตัวแปร y เราต้องคูณสมการแรกด้วย a_{22} และคูณสมการที่สองด้วย a_{12} แล้วลบกันผลลัพธ์คือ

$$(a_{11}a_{22} - a_{21}a_{12})\,x = a_{22}b_1 - a_{12}b_2$$

ซึ่งเราจะได้ค่าของตัวแปรเป็น

$$x = \frac{a_{22}b_1 - a_{12}b_2}{(a_{11}a_{22} - a_{21}a_{12})}$$

ในทำนองเดียวกัน เมื่อพยายามแก้หาตัวแปร y แทนเราจะได้สมการ

$$y = \frac{a_{11}b_2 - a_{21}b_1}{(a_{11}a_{22} - a_{21}a_{12})}$$

ซึ่งสังเกตได้ว่ามีค่าคงที่ตัวหนึ่งที่ปรากฏตัวขึ้นบ่อยในการแก้สมการ เราจะนิยามมันว่าดีเทอมิแนนท์ (determinant) ของแมทริกซ์

$$\hat{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \to \det \hat{A} \equiv (a_{11}a_{22} - a_{21}a_{12})$$

ถ้าเราพยายามเขียนสมการตั้งต้นในรูปแบบของตัวแทนแมทริกซ์ จะได้ว่า

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \to \hat{A}|x\rangle = |b\rangle$$

และคำตอบสามารถเขียนได้เป็น

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\det \hat{A}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \to |x\rangle = \hat{A}^{-1}|b\rangle$$

ซึ่งเราจะเห็นได้ว่าเราสามารถเขียนอินเวอร์สแมทริกซ์ได้เป็น

$$\hat{A}^{-1} = \frac{1}{\det \hat{A}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

แมทริกซ์ใดๆ จะไม่มีอินเวอร์สก็ต่อเมื่อ

$$\det \hat{A} = 0$$

แมทริกซ์ที่หาอินเวอร์สไม่ได้จะเรียกว่า non-invertible matrix หรือ singular matrix
ในกรณีของสามมิติและมิติที่สูงขึ้นเราจะศึกษาการหาอินเวอร์สและดีเทอมิแนนท์ในกรณีทั่วไปในบทถัดไป

ตัวอย่าง จงหาดีเทอมิแนนท์ของแมทริกซ์การหมุน พร้อมทั้งหาอินเวอร์สของแมทริกซ์การหมุนนี้

$$\hat{T} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

ตัวอย่าง จงหาดีเทอมิแนนท์ของการสะท้อนรอบแกน $x = y$ พร้อมทั้งหาอินเวอร์สของแมทริกซ์

$$\hat{T} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

เพื่อที่จะศึกษาจาโคเบียน เราพิจารณาการเปลี่ยนพิกัด (coordinate) $(x, y) \to (u, v)$ ในสองมิติ

$$x = f(u, v), \qquad y = g(u, v)$$

เราจะพบว่า differential สามารถเขียนเป็น

$$dx = \frac{\partial f}{\partial u} du + \frac{\partial f}{\partial v} dv, \qquad dy = \frac{\partial g}{\partial u} du + \frac{\partial g}{\partial v} dv$$

ถ้าพิจารณา primary curve แรกโดยการเพิ่ม u และคงค่าของ v ไว้ เราจะได้

$$d\vec{l}_1 = dx_u \hat{e}_x + dy_u \hat{e}_y = \frac{\partial f}{\partial u} du\ \hat{e}_x + \frac{\partial g}{\partial u} du\ \hat{e}_y$$

และ primary curve ที่สองซึ่งเกิดจากการเพิ่ม v และคงค่าของ u ไว้ เราจะได้

$$d\vec{l}_2 = dx_v \hat{e}_x + dy_v \hat{e}_y = \frac{\partial f}{\partial v} dv\ \hat{e}_x + \frac{\partial g}{\partial v} dv\ \hat{e}_y$$

เมื่อทำการหา cross product จะได้

$$d\vec{l}_1 \times d\vec{l}_2 = \det \begin{pmatrix} \hat{e}_x & \hat{e}_y & \hat{e}_z \\ \frac{\partial f}{\partial u} du & \frac{\partial g}{\partial u} du & 0 \\ \frac{\partial f}{\partial v} dv & \frac{\partial g}{\partial v} dv & 0 \end{pmatrix} = \det \begin{pmatrix} \frac{\partial f}{\partial u} & \frac{\partial g}{\partial u} \\ \frac{\partial f}{\partial v} & \frac{\partial g}{\partial v} \end{pmatrix} dudv\ \hat{e}_z$$

ดังนั้นพื้นที่ที่เกิดจาก primary curves สองเส้นนี้จะกลายเป็น

$$\boxed{da = \left| d\vec{l}_1 \times d\vec{l}_2 \right| = \det \begin{pmatrix} \frac{\partial f}{\partial u} & \frac{\partial g}{\partial u} \\ \frac{\partial f}{\partial v} & \frac{\partial g}{\partial v} \end{pmatrix} dudv \equiv \begin{vmatrix} \frac{\partial f}{\partial u} & \frac{\partial g}{\partial u} \\ \frac{\partial f}{\partial v} & \frac{\partial g}{\partial v} \end{vmatrix} dudv}$$

ซึ่งแมทริกซ์ที่ใช้ทำการหาดีเทอมิแนนท์เราเรียกว่า**จาโคเบียนแมทริกซ์ (Jacobian matrix)**

และค่าดีเทอมิแนนท์เราเรียกว่า**จาโคเบียน (Jacobian)**

$$J=\begin{vmatrix}\frac{\partial f}{\partial u} & \frac{\partial g}{\partial u}\\ \frac{\partial f}{\partial v} & \frac{\partial g}{\partial v}\end{vmatrix}$$

ในกรณีของสามมิติเราสามารถขยายผลนี้ได้โดยการมองเป็นปริมาตร เริ่มจากการเปลี่ยนระบบพิกัด $(x,y,z)\rightarrow(u,v,w)$

$$x=f(u,v,w),\qquad y=g(u,v,w),\qquad z=h(u,v,w)$$

Primary curves ทั้งสามเส้นจะเขียนได้เป็น

$$d\vec{l}_1=dx_u\hat{e}_x+dy_u\hat{e}_y+dz_u\hat{e}_z=\frac{\partial f}{\partial u}du\;\hat{e}_x+\frac{\partial g}{\partial u}du\;\hat{e}_y+\frac{\partial h}{\partial u}du\;\hat{e}_z$$

$$d\vec{l}_2=dx_v\hat{e}_x+dy_v\hat{e}_y+dz_v\hat{e}_z=\frac{\partial f}{\partial v}dv\;\hat{e}_x+\frac{\partial g}{\partial v}dv\;\hat{e}_y+\frac{\partial h}{\partial v}dv\;\hat{e}_z$$

$$d\vec{l}_3=dx_w\hat{e}_x+dy_w\hat{e}_y+dz_w\hat{e}_z=\frac{\partial f}{\partial w}dw\;\hat{e}_x+\frac{\partial g}{\partial w}dw\;\hat{e}_y+\frac{\partial h}{\partial w}dw\;\hat{e}_z$$

ดังนั้นปริมาตรหาได้จาก

$$dV=\left|d\vec{l}_1\cdot\left(d\vec{l}_2\times d\vec{l}_3\right)\right|=\det\begin{pmatrix}\frac{\partial f}{\partial u}du & \frac{\partial g}{\partial u}du & \frac{\partial h}{\partial u}du\\ \frac{\partial f}{\partial v}dv & \frac{\partial g}{\partial v}dv & \frac{\partial h}{\partial v}dv\\ \frac{\partial f}{\partial w}dw & \frac{\partial g}{\partial w}dw & \frac{\partial h}{\partial w}dw\end{pmatrix}=\begin{vmatrix}\frac{\partial f}{\partial u} & \frac{\partial g}{\partial u} & \frac{\partial h}{\partial u}\\ \frac{\partial f}{\partial v} & \frac{\partial g}{\partial v} & \frac{\partial h}{\partial v}\\ \frac{\partial f}{\partial w} & \frac{\partial g}{\partial w} & \frac{\partial h}{\partial w}\end{vmatrix}dudvdw$$

ซึ่งจาโคเบียนจะเป็น

$$J=\begin{vmatrix}\frac{\partial f}{\partial u} & \frac{\partial g}{\partial u} & \frac{\partial h}{\partial u}\\ \frac{\partial f}{\partial v} & \frac{\partial g}{\partial v} & \frac{\partial h}{\partial v}\\ \frac{\partial f}{\partial w} & \frac{\partial g}{\partial w} & \frac{\partial h}{\partial w}\end{vmatrix}$$

ตัวอย่าง จงพิจารณาพื้นที่ในพิกัดเชิงขั้วที่มีการแปลงระหว่างพิกัดเชิงขั้วและพิกัดคาร์ทีเซียนเป็น

$$x = f(r,\theta) = r\cos\theta\,, \qquad y = g(r,\theta) = r\sin\theta$$

ตัวอย่าง จงพิจารณาพื้นที่ในพิกัดทรงกลมที่มีการแปลงระหว่างพิกัดทรงกลมและพิกัดคาร์ทีเซียนเป็น

$$x = f(r,\theta,\phi) = r\sin\theta\cos\phi$$
$$y = g(r,\theta,\phi) = r\sin\theta\sin\phi$$
$$z = h(r,\theta,\phi) = r\cos\theta$$

แบบฝึกหัด 2

1) Euler angles (มุมออยเลอร์)

1.1) จงแสดงว่าการหมุนแกนในระนาบ xy ด้วยมุม φ ซึ่งมีการเปลี่ยนเวกเตอร์ฐาน

$\{|e_x\rangle, |e_y\rangle, |e_z\rangle\} \to \{|e'_x\rangle, |e'_y\rangle, |e'_z\rangle\}$ เขียนได้เป็น

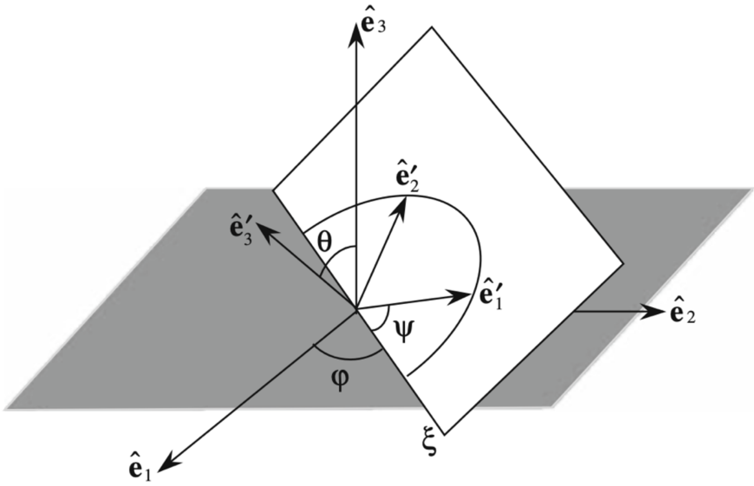
$$\hat{R}_1|e_x\rangle \to |e'_x\rangle = \cos\varphi\,|e_x\rangle + \sin\varphi\,|e_y\rangle$$

$$\hat{R}_1|e_y\rangle \to |e'_y\rangle = -\sin\varphi\,|e_x\rangle + \cos\varphi\,|e_y\rangle$$

$$\hat{R}_1|e_z\rangle \to |e'_z\rangle = |e_z\rangle$$

จะมีตัวแทนแมทริกซ์เป็น

$$\tilde{R}_1 = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$



1.2) ถ้าพิจารณาการหมุนแกนอีกครั้งในระนาบ $y'z'$ ด้วยมุม θ ตามภาพจะมีตัวแทนแมทริกซ์เป็น

$$\tilde{R}_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}$$

หลังจากนั้นถ้าทำการหมุนแกนเป็นครั้งสุดท้ายในระนาบ $x''y''$ ด้วยมุม ψ ตามภาพจะมีตัวแทนแมทริกซ์เป็น

$$\tilde{R}_3 = \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

จงแสดงว่าผลคูณของการหมุนทั้งหมดนี้คือ

$$\tilde{R} = \tilde{R}_3\tilde{R}_2\tilde{R}_1$$

$$= \begin{pmatrix} \cos\psi\cos\varphi - \sin\psi\cos\theta\sin\varphi & -\cos\psi\sin\varphi - \sin\psi\cos\theta\cos\varphi & \sin\psi\sin\theta \\ \sin\psi\cos\varphi + \cos\psi\cos\theta\sin\varphi & -\sin\psi\sin\varphi + \cos\psi\cos\theta\cos\varphi & -\cos\psi\sin\theta \\ \sin\theta\sin\varphi & \sin\theta\cos\varphi & \cos\theta \end{pmatrix}$$

ซึ่งมุม (ψ, φ, ϕ) เป็นตัวแทนของมุมออยเลอร์ (Euler angles)

2) Rotation and Orthogonality condition (การหมุนและเงื่อนไขการตั้งฉาก)

2.1) จากการหมุนในสองมิติ

$$\hat{T}_2 = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

จงแสดงว่าการหาขนาดของเวกเตอร์ผ่านเมทริกแมทริกซ์เอกลักษณ์ไม่เปลี่ยนภายใต้การหมุน นั่นคือ ถ้า

$$|v'\rangle = \hat{T}_2|v\rangle \rightarrow \begin{pmatrix} v_1' \\ v_2' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

แล้วเงื่อนไขตั้งฉากจะเป็นจริง

$$\langle v'|\mathbb{I}|v'\rangle = \langle v|\mathbb{I}|v\rangle$$

เมื่อ

$$\mathbb{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

2.2) จากการหมุนด้วยมุมออยเลอร์ในสามมิติ (ข้อที่แล้ว) จงแสดงว่าการหาขนาดของเวกเตอร์ผ่านเมทริกแมทริกซ์ไม่เปลี่ยนภายใต้การหมุน นั่นคือถ้า

$$|v'\rangle = \hat{R}|v\rangle \rightarrow \begin{pmatrix} v_1' \\ v_2' \\ v_3' \end{pmatrix}$$

$$= \begin{pmatrix} \cos\psi\cos\varphi - \sin\psi\cos\theta\sin\varphi & -\cos\psi\sin\varphi - \sin\psi\cos\theta\cos\varphi & \sin\psi\sin\theta \\ \sin\psi\cos\varphi + \cos\psi\cos\theta\sin\varphi & -\sin\psi\sin\varphi + \cos\psi\cos\theta\cos\varphi & -\cos\psi\sin\theta \\ \sin\theta\sin\varphi & \sin\theta\cos\varphi & \cos\theta \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

แล้วเงื่อนไขตั้งฉากจะเป็นจริง

$$\langle v'|\mathbb{I}|v'\rangle = \langle v|\mathbb{I}|v\rangle$$

เมื่อ

$$\mathbb{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

<u>3) Special Relativity (สัมพัทธภาพพิเศษ)</u>

3.1) กำหนดให้การแปลงกรอบอ้างอิงในสัมพัทธภาพพิเศษเป็นการแปลงแบบลอเรนซ์ (Lorentz transformation)

$$ct \rightarrow ct' = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} ct + \frac{\frac{v}{c}}{\sqrt{1 - \frac{v^2}{c^2}}} x$$

$$x \rightarrow x' = \frac{\frac{v}{c}}{\sqrt{1 - \frac{v^2}{c^2}}} ct + \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} x$$

โดยที่ c เป็นความเร็วแสงและ v เป็นความเร็วของกรอบอ้างอิงใหม่เทียบกับกรอบอ้างอิงเดิม ถ้ากำหนดให้เวกเตอร์เขียนได้จากตำแหน่งและเวลาดังนี้

$$|v\rangle = \begin{pmatrix} ct \\ x \end{pmatrix}$$

ถ้าพิจารณาเวกเตอร์ฐานเป็น $|\mathrm{e}_1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, |\mathrm{e}_2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

จงหาแมทริกซ์ตัวแทนของการแปลงลอเรนซ์

3.2) ถ้ากำหนดให้แมทริกซ์ตัวแทนของการแปลงลอเรนซ์เป็น Λ จงแสดงว่าการหาขนาดของเวกเตอร์ผ่านเมทริกแมทริกซ์ $\hat{G}$ ไม่เปลี่ยนภายใต้การหมุน นั่นคือถ้า

$$|v'\rangle = \hat{\Lambda}|v\rangle$$

แล้วเงื่อนไขตั้งฉากจะเป็นจริง

$$\langle v'|\hat{G}|v'\rangle = \langle v|\hat{G}|v\rangle$$

เมื่อ

$$\hat{G} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

4) Matrix Operation Identity (เอกลักษณ์ของตัวดำเนินการบนแมทริกซ์)

กำหนดให้ $\hat{A} = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}$ และ $\hat{B} = \begin{pmatrix} a_2 & b_2 \\ c_2 & d_2 \end{pmatrix}$ จงทำการคำนวณเพื่อพิสูจน์อย่างชัดแจ้ง (เขียนกระจายในรูปองค์ประกอบ) ว่า

4.1) $(\hat{A}\hat{B})^T = \hat{B}^T\hat{A}^T$

4.2) $(\hat{A} + \hat{B})^T = \hat{A}^T + \hat{B}^T$

4.3) $\det(\hat{A}\hat{B}) = \det(\hat{B}\hat{A})$

4.4) $\det(\hat{A}\hat{B}) = \det(\hat{A})\det(\hat{B})$

4.5) $\det(\hat{A}^{-1}) = \frac{1}{\det(\hat{A})}$

5) Inverse Matrix (การหาอินเวอร์ส)

จงหา inverse ของแมทริกซ์ต่อไปนี้

5.1) $\tilde{G} = \begin{pmatrix} 2 & -1 \\ -1 & 3 \end{pmatrix}$

5.2) $\tilde{G} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

5.3) $\tilde{G} = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$

<u>6) Jacobian (การหาจาโคเบียน)</u>

6.1) กำหนดให้การแปลงระบบพิกัดจากพิกัดคาร์ทิเซียนเป็นระบบพิกัดทรงกระบอก (cylindrical coordinates) นั่นคือ $(x, y, z) \to (\rho, \theta, z)$ มีการแปลงดังนี้

$$x = \rho\cos\theta$$

$$y = \rho\sin\theta$$

$$z = z$$

จงใช้จาโคเบียนของการแปลงพิกัดนี้เพื่อหาปริมาตร dV ในพิกัดทรงกระบอก

6.2) กำหนดให้การแปลงระบบพิกัดจากพิกัดคาร์ทิเซียนเป็นระบบพิกัดทรงรีกระบอก (elliptical cylindrical coordinates) นั่นคือ $(x, y, z) \to (u, \theta, z)$ มีการแปลงดังนี้

$$x = a\cosh u\cos\theta$$

$$y = a\sinh u\sin\theta$$

$$z = z$$

โดยที่ a เป็นค่าคงที่ จงใช้จาโคเบียนของการแปลงพิกัดนี้เพื่อหาปริมาตร dV ในพิกัดทรงรีกระบอก

6.3) กำหนดให้การแปลงระบบพิกัดจากพิกัดคาร์ทิเซียนเป็นระบบทรงกลมแบบขยายออก (prolate spherical coordinates) นั่นคือ $(x, y, z) \to (u, \theta, \phi)$ มีการแปลงดังนี้

$$x = a\sinh u\sin\theta\cos\phi$$

$$y = a\sinh u\sin\theta\sin\phi$$

$$z = a\cosh u\cos\theta$$

โดยที่ a เป็นค่าคงที่ จงใช้จาโคเบียนของการแปลงพิกัดนี้เพื่อหาปริมาตร dV ในพิกัดทรงกลมแบบขยายออก

Chapter 3: Application

3.1) การแก้ระบบสมการเชิงเส้น (Systems of linear equations)

ในบทนี้เราจะใช้ความรู้เกี่ยวกับตัวดำเนินการเชิงเส้นและพีชคณิตเชิงเส้นมาเพื่อแก้ปัญหาระบบสมการเชิงเส้น โดยเราจะเริ่มจากการทำความเข้าใจระบบสมการเชิงเส้นก่อน

สมการเชิงเส้นคือสมการที่มีตัวแปรไม่ทราบค่า (unknown variables) อยู่ในลักษณะของผลรวมเชิงเส้น ตัวอย่างเช่น

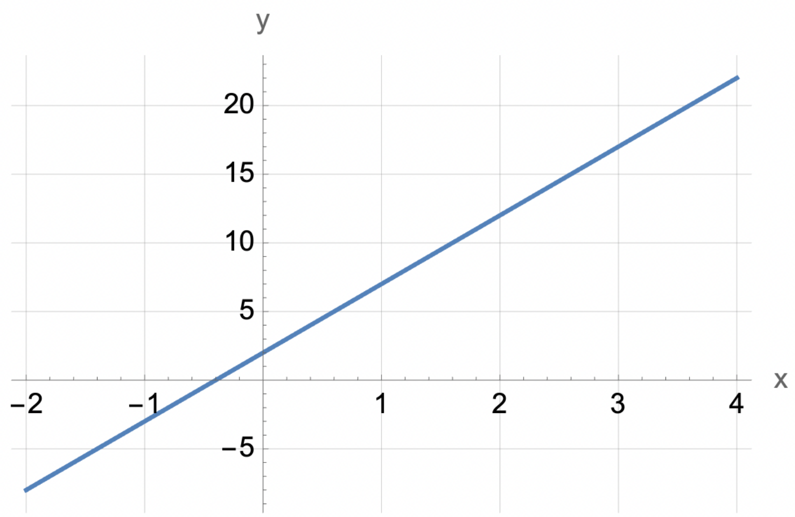
$$a_{11}x + a_{12}y = b_1$$

เป็นสมการเชิงเส้นของตัวแปร (x, y) โดยที่มี a_{11}, a_{12}, b_1 เป็นค่าคงที่

ระบบสมการเชิงเส้นคือสมการเชิงเส้นหลายสมการ โดยทั่วไปเราต้องการจำนวนสมการเชิงเส้นเท่ากับจำนวนตัวแปรเพื่อที่จะมีโอกาสในการแก้ระบบสมการนี้ได้

เราสามารถตีความสมการเชิงเส้นเป็นเส้นตรงในปริภูมิเวกเตอร์ได้ตัวอย่างเช่น

สมการ $-10x + 2y = 4$ สอดคล้องกับกราฟ

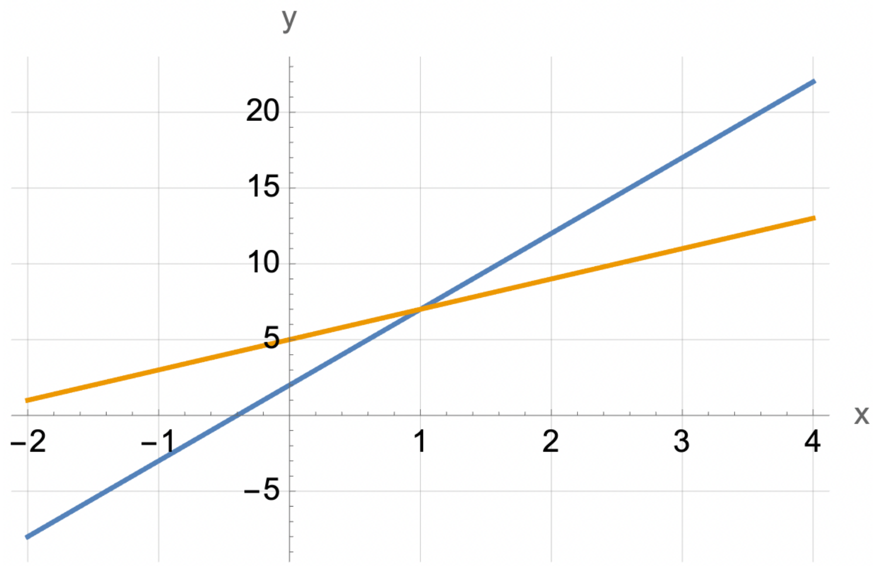


โดยกราฟนี้มีความหมายว่าถ้าต้องการให้จุด (x, y) เป็นคำตอบของสมการดังกล่าวจุดนี้ต้องอยู่บนเส้นกราฟนี้

ดังนั้นถ้าเรามีสมการสองสมการที่ต้องการให้เป็นจริงทั้งคู่ คำตอบที่ได้จากสองสมการนี้ต้องอยู่บนเส้นทั้งสองด้วย ตัวอย่างเช่นถ้าสมการทั้งสองสมการเป็น

$-10x + 2y = 4$ และ $-2x + y = 5$

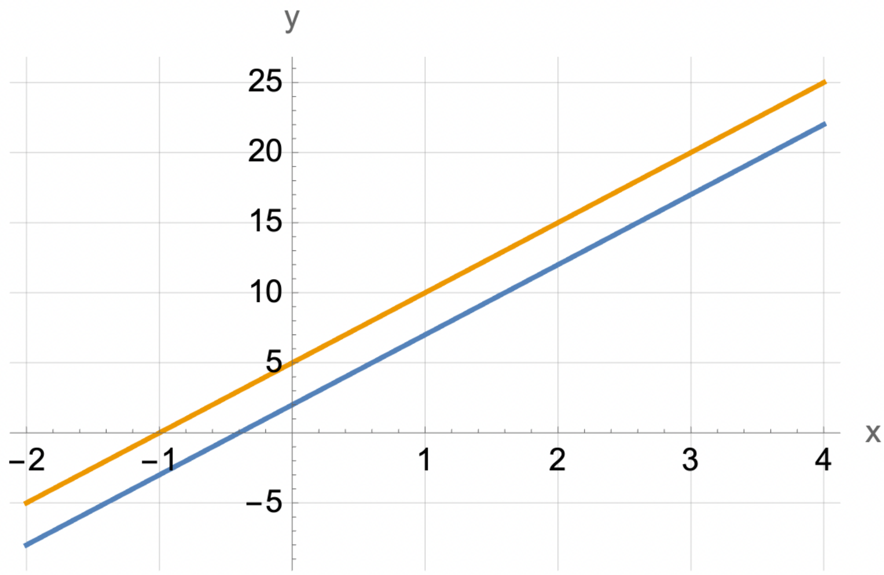
เราจะได้กราฟเป็น



คำตอบของสมการทั้งสองนี้คือจุดที่อยู่บนเส้นตรงทั้งสองนั่นคือจุด $(1,7)$

อย่างไรก็ดีถ้าเส้นตรงทั้งสองนี้ขนานกันเราจะได้อีกสองกรณี เช่นถ้าสมการทั้งสองเป็น

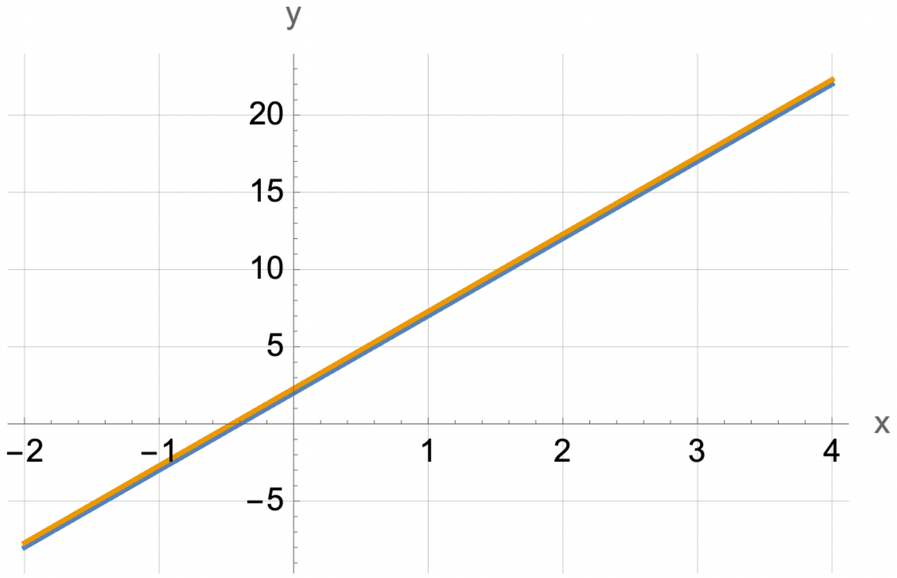
$-10x + 2y = 4$ และ $-5x + y = 5$ เราจะได้กราฟเป็น



ซึ่งจะทำให้ระบบสมการทั้งสองนี้ไม่มีคำตอบ เพราะไม่มีจุดใดเลยที่อยู่บนเส้นทั้งสอง

ในขณะที่ถ้าสมการทั้งสองเป็น

$-10x + 2y = 4$ และ $-5x + y = 2$ เราจะได้เส้นกราฟสองเส้นที่ทับกันพอดี



ซึ่งจะทำให้คำตอบของสมการคือทุกจุดบนเส้นตรงนี้ หรืออีกความหมายคือสมการทั้งสองนี้มีความหมายเดียวกัน

เทคนิคในการแก้ระบบสมการเชิงเส้นโดยใช้วิธีการของแมทริกซ์นั้นเริ่มจากการสังเกตระบบสมการเชิงเส้นสองตัวแปร

$$R_1: a_{11}x + a_{12}y = b_1$$
$$R_2: a_{21}x + a_{22}y = b_2$$

โดยสัญลักษณ์ R_1, R_2 คือการบอกถึงแถวของสมการ
เราจะเริ่มจากพิจารณาการเปลี่ยนแปลงของระบบสมการที่ทำให้คำตอบของระบบสมการนี้ไม่เปลี่ยนแปลง นั่นคือได้คำตอบเดิม โดยเราพบว่ามีกระบวนการต่อไปนี้ที่ทำให้ไม่เปลี่ยนคำตอบของระบบสมการ

1) การคูณสมการด้วยค่าคงที่ $R_1 \to kR_1$

$$R_1: a_{11}x + a_{12}y = b_1 \to \boldsymbol{kR_1: ka_{11}x + ka_{12}y = kb_1}$$
$$R_2: a_{21}x + a_{22}y = b_2 \to \boldsymbol{R_2: a_{21}x + a_{22}y = b_2}$$

2) การสลับแถว $\boldsymbol{R_i \leftrightarrow R_j}$

$$R_1: a_{11}x + a_{12}y = b_1 \to \boldsymbol{R_2: a_{21}x + a_{22}y = b_2}$$
$$R_2: a_{21}x + a_{22}y = b_2 \to \boldsymbol{R_1: a_{11}x + a_{12}y = b_1}$$

3) การคูณแถวแล้วค่าคงที่แล้วนำไปบวกกับอีกแถวหนึ่ง $\boldsymbol{R_2 \to R_2 + kR_1}$

$$R_1: a_{11}x + a_{12}y = b_1 \to \boldsymbol{R_1: a_{11}x + a_{12}y = b_1}$$
$$R_2: a_{21}x + a_{22}y = b_2 \to \boldsymbol{R_2 + kR_1: (a_{21} + ka_{11})x + (a_{22} + ka_{12})y = b_2 + kb_1}$$

โดยสังเกตว่าการเปลี่ยนแปลงกับระบบสมการนี้มีคำตอบเหมือนเดิมทุกประการ หลังจากนี้เราจะไม่เขียนระบบสมการด้วยการระบุแถวและเครื่องหมายสมการอีกต่อไป เพราะการเปลี่ยนแปลงของระบบสมการเหล่านี้ทำกับค่าคงที่เท่านั้น นั่นคือ

$$R_1: a_{11}x + a_{12}y = b_1$$
$$R_2: a_{21}x + a_{22}y = b_2$$

มีค่าเท่ากับการเขียน**แมทริกซ์แต่งเติม (augmented matrix)**

$$\left(\begin{array}{cc|c} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \end{array}\right)$$

ซึ่งเรากระบวนการเปลี่ยนระบบสมการด้านบนจะถูกเรียกว่ากระบวนการแถว (row operations) คือกระบวนการที่ไม่เปลี่ยนคำตอบของระบบสมการ ซึ่งประกอบไปด้วย

1) การคูณด้วยค่าคงที่

$$\left(\begin{array}{cc|c} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \end{array}\right) \xrightarrow{R_1 \to kR_1} \left(\begin{array}{cc|c} ka_{11} & ka_{12} & kb_1 \\ a_{21} & a_{22} & b_2 \end{array}\right)$$

2) การสลับแถว

$$\left(\begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{matrix}\middle| \begin{matrix} b_1 \\ b_2 \end{matrix}\right) \xrightarrow{R_1 \leftrightarrow R_2} \left(\begin{matrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{matrix}\middle| \begin{matrix} b_2 \\ b_1 \end{matrix}\right)$$

3) การคูณแถวแล้วค่าคงที่แล้วนำไปบวกกับอีกแถวหนึ่ง

$$\left(\begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{matrix}\middle| \begin{matrix} b_1 \\ b_2 \end{matrix}\right) \xrightarrow{R_2 \to R_2 + kR_1} \left(\begin{matrix} a_{11} & a_{12} \\ a_{21} + ka_{11} & a_{22} + ka_{12} \end{matrix}\middle| \begin{matrix} b_1 \\ b_2 + kb_1 \end{matrix}\right)$$

กระบวนการทั้งสามที่นำไปใช้เพื่อการแก้สมการจะถูกเรียกว่า**กระบวนการตัดออกของเกาส์ (Gaussian elimination methods)** โดยเป้าหมายหลักคือต้องพยายามทำให้ฝั่งซ้ายของเส้นคั่นมีลักษณะเป็นแมทริกซ์เอกลักษณ์นั่นคือ ถ้าแมทริกซ์ตกแต่งตั้งต้นสามารถทำให้อยู่ในรูป

$$\left(\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}\middle| \begin{matrix} b'_1 \\ b'_2 \end{matrix}\right)$$

ถ้าเราทำการแปลงจากแมทริกซ์ตกแต่งไปเป็นระบบสมการเราจะบอกได้ทันทีว่าคำตอบคือ

$$1 \cdot x + 0 \cdot y = \boldsymbol{x} = \boldsymbol{b'_1}$$

$$0 \cdot x + 1 \cdot y = \boldsymbol{y} = \boldsymbol{b'_2}$$

อย่างไรก็ดีถ้าการทำให้ฝั่งซ้ายของเส้นคั่นเป็นแมทริกซ์เอกลักษณ์ใช้ขั้นตอนการคำนวณที่มากเกินความจำเป็น ในหลายครั้งเราจะทำเพียงแค่ครึ่งเดียว นั่นคือเราจะพยายามทำให้อยู่ในรูปของแมทริกซ์ขั้นบันไดลดรูป (reduced row echelon form) เช่น

$$\left(\begin{matrix} 1 & a \\ 0 & 1 \end{matrix}\middle| \begin{matrix} b'_1 \\ b'_2 \end{matrix}\right)$$

ซึ่งเราจะได้คำตอบตัวแปรหนึ่งทันที

$$y = b'_2$$

หลังจากนั้นตัวแปรที่เหลือสามารถแทนค่าเพื่อได้คำตอบกลับมาได้

$$1 \cdot x + a \cdot y = x + ab'_2 = b'_1$$

$$x = b'_1 - ab'_2$$

ตัวอย่าง จงใช้กระบวนการตัดออกของเกาส์เพื่อลดระบบสมการเชิงเส้น

$$2x + y - z = 8$$

$$-3x - y + 2z = -11$$

$$-2x + y + 2z = -3$$

ให้อยู่ในรูป reduced row echelon form

$$\left(\begin{array}{ccc|c} 1 & 1/2 & -1/2 & 4 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & -1 \end{array}\right)$$

พร้อมทั้งหาคำตอบของ (x, y, z)

ตัวอย่าง จงใช้กระบวนการตัดออกของเกาส์เพื่อลดระบบสมการเชิงเส้น

$$x - 3y + 4z = 3$$

$$2x - 5y + 6z = 6$$

$$-3x + 3y + 4z = 6$$

ให้อยู่ในรูป reduced row echelon form

$$\left(\begin{array}{ccc|c} 1 & -3 & 4 & 3 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 1 & 15/4 \end{array}\right)$$

พร้อมทั้งหาคำตอบของ (x, y, z)

3.2) การหาอินเวอร์สแมทริกซ์โดยใช้ตัวดำเนินการแถว (Inverse Matrix from Row Operations)

อีกการประยุกต์ใช้ตัวดำเนินการแถวหรือกระบวนการตัดออกของเกาส์คือการทำเพื่อหาอินเวอร์สแมทริกซ์ (Inverse Matrix) เราจะเริ่มพิจารณาการคูณกันของแมทริกซ์สี่เหลี่ยมขนาด 3x3

$$\hat{A}\hat{B}=\hat{C}\rightarrow\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{pmatrix}\begin{pmatrix}b_{11}&b_{12}&b_{13}\\b_{21}&b_{22}&b_{23}\\b_{31}&b_{32}&b_{33}\end{pmatrix}=\begin{pmatrix}c_{11}&c_{12}&c_{13}\\c_{21}&c_{22}&c_{23}\\c_{31}&c_{32}&c_{33}\end{pmatrix}$$

โดยเราจะมองปัญหานี้เป็นการสมการเวกเตอร์สามสมการนั่นคือ

$$\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{pmatrix}\begin{pmatrix}b_{11}\\b_{21}\\b_{31}\end{pmatrix}=\begin{pmatrix}c_{11}\\c_{21}\\c_{31}\end{pmatrix}$$

$$\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{pmatrix}\begin{pmatrix}b_{12}\\b_{22}\\b_{32}\end{pmatrix}=\begin{pmatrix}c_{12}\\c_{22}\\c_{32}\end{pmatrix}$$

$$\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{pmatrix}\begin{pmatrix}b_{13}\\b_{23}\\b_{33}\end{pmatrix}=\begin{pmatrix}c_{13}\\c_{23}\\c_{33}\end{pmatrix}$$

ซึ่งคำตอบของสมการทั้งสามนี้ไม่เปลี่ยนภายใต้ตัวดำเนินการแถว ดังนั้นการเราจึงสามารถใช้ตัวดำเนินการแถวกับแมทริกซ์ตกแต่งต่อไปนี้ได้โดยไม่เปลี่ยนผลลัพธ์ของการคูณ

$$\hat{A}\hat{B}=\hat{C}\rightarrow\left(\begin{array}{ccc|ccc}a_{11}&a_{12}&a_{13}&c_{11}&c_{12}&c_{13}\\a_{21}&a_{22}&a_{23}&c_{21}&c_{22}&c_{23}\\a_{31}&a_{32}&a_{33}&c_{31}&c_{32}&c_{33}\end{array}\right)$$

จากความจริงข้อนี้เราสามารถประยุกต์ใช้กับการหาอินเวอร์สของแมทริกซ์ได้โดยการพิจารณาสมบัติของอินเวอร์สในรูปของแมทริกซ์ตกแต่ง

$$\hat{A}\hat{A}^{-1}=\hat{\mathbb{I}}\rightarrow\left(\begin{array}{ccc|ccc}a_{11}&a_{12}&a_{13}&1&0&0\\a_{21}&a_{22}&a_{23}&0&1&0\\a_{31}&a_{32}&a_{33}&0&0&1\end{array}\right)$$

ดังนั้นถ้าเราสามารถใช้ตัวดำเนินการแถวพยายามจัดให้ฝั่งซ้ายของเส้นคั่นเป็นแมทริกซ์เอกลักษณ์นั่นคือ

$$\left(\begin{array}{ccc|ccc}a_{11}&a_{12}&a_{13}&1&0&0\\a_{21}&a_{22}&a_{23}&0&1&0\\a_{31}&a_{32}&a_{33}&0&0&1\end{array}\right)\xrightarrow{Row\ operations}\left(\begin{array}{ccc|ccc}1&0&0&A'_{11}&A'_{12}&A'_{13}\\0&1&0&A'_{21}&A'_{22}&A'_{23}\\0&0&1&A'_{31}&A'_{32}&A'_{33}\end{array}\right)$$

เราจะแปลผลของการคูณแมทริกซ์นี้เป็น

$$\hat{A}\hat{A}^{-1}=\hat{\mathbb{I}}\rightarrow\hat{\mathbb{I}}\hat{A}^{-1}=\hat{A}^{-1}$$

ซึ่งมีความหมายว่า

$$\hat{A}^{-1} = \begin{pmatrix} A'_{11} & A'_{12} & A'_{13} \\ A'_{21} & A'_{22} & A'_{23} \\ A'_{31} & A'_{32} & A'_{33} \end{pmatrix}$$

ตัวอย่าง จงหาอินเวอร์สของ

$$\hat{A} = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

โดยการใช้ตัวดำเนินการแถวกับแมทริกซ์ตกแต่ง

$$\left(\begin{array}{ccc|ccc} 2 & -1 & 0 & 1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 1 & 0 \\ 0 & -1 & 2 & 0 & 0 & 1 \end{array}\right)$$

ตัวอย่าง จงหาอินเวอร์สของ

$$\hat{A} = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

โดยการใช้ตัวดำเนินการแถวกับแมทริกซ์ตกแต่ง

$$\left(\begin{array}{ccc|ccc} 2 & -1 & 0 & 1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 1 & 0 \\ 0 & -1 & 2 & 0 & 0 & 1 \end{array}\right)$$

ตัวอย่าง จงหาอินเวอร์สของ

$$\hat{A} = \begin{pmatrix} 6 & 2 & 3 \\ 3 & 1 & 1 \\ 10 & 3 & 4 \end{pmatrix}$$

โดยการใช้ตัวดำเนินการแถวกับแมทริกซ์ตกแต่ง

$$\left(\begin{array}{ccc|ccc} 6 & 2 & 3 & 1 & 0 & 0 \\ 3 & 1 & 1 & 0 & 1 & 0 \\ 10 & 3 & 4 & 0 & 0 & 1 \end{array}\right)$$

3.3) การหาดีเทอมิแนนท์โดยตัวดำเนินการแถว (Determinant from Row Operations)

อีกการประยุกต์ใช้ของตัวดำเนินการแถวคือการหาดีเทอมิแนนท์ของแมทริกซ์ใดๆ เราเริ่มจากการพิจารณาผลของตัวดำเนินการแถวบนดีเทอมิแนนท์ของแมทริกซ์ขนาด 2x2

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

1) การคูณด้วยค่าคงที่

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \xrightarrow{R_1 \to kR_1} \begin{pmatrix} ka_{11} & ka_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

$$\boxed{\det \begin{pmatrix} ka_{11} & ka_{12} \\ a_{21} & a_{22} \end{pmatrix} = k \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}}$$

2) การสลับแถว

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{pmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{pmatrix}$$

$$\boxed{\det \begin{pmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{pmatrix} = (-1) \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}}$$

3) การคูณแถวแล้วค่าคงที่แล้วนำไปบวกกับอีกแถวหนึ่ง

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \xrightarrow{R_2 \to R_2 + kR_1} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} + ka_{11} & a_{22} + ka_{12} \end{pmatrix}$$

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} + ka_{11} & a_{22} + ka_{12} \end{pmatrix} = a_{11}(a_{22} + ka_{12}) - (a_{21} + ka_{11})a_{12} = a_{11}a_{22} - a_{21}a_{12}$$

$$\boxed{\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} + ka_{11} & a_{22} + ka_{12} \end{pmatrix} = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}}$$

ด้วยความสัมพันธ์ของค่าดีเทอมิแนนท์จากตัวดำเนินการแถว เราสามารถใช้ตัวดำเนินการแถวเพื่อปรับให้เป็นแมทริกซ์ขั้นบันไดลดรูปซึ่งมีค่าดีเทอมิแนนท์เป็น 1 และหาดีเทอมิแนนท์จากความสัมพันธ์

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \xrightarrow{Row\ operations} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\det\begin{pmatrix} 1 & A & B \\ 0 & 1 & C \\ 0 & 0 & 1 \end{pmatrix} = 1 = d \cdot \det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

โดยที่ค่า d เกิดจากผลคูณของค่าคงที่ที่เกิดจากการคูณแถวด้วยค่าคงที่ และ (-1) จากการสลับแถวตามตัวดำเนินการแถวที่ได้ทำมาทั้งหมด และค่าดีเทอมิแนนท์จะได้เป็น

$$\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \frac{1}{d}$$

ตัวอย่าง จงคำนวณดีเทอมิแนนท์ของ

$$\det\begin{pmatrix} 1 & 1 & 2 \\ 3 & 1 & 2 \\ 3 & 1 & 1 \end{pmatrix}$$

โดยการใช้ตัวดำเนินการแถว

ตัวอย่าง จงคำนวณดีเทอมิแนนท์ของ

$$\det\begin{pmatrix} 1 & 4 & 2 & 3 \\ 0 & 1 & 4 & 4 \\ -1 & 0 & 1 & 0 \\ 2 & 0 & 4 & 1 \end{pmatrix}$$

โดยการใช้ตัวดำเนินการแถว

ตัวอย่าง จงคำนวณดีเทอมิแนนท์ของ

$$\det\begin{pmatrix} 4 & 3 & 2 & 2 \\ 0 & 1 & -3 & 3 \\ 0 & -1 & 3 & 3 \\ 0 & 3 & 1 & 1 \end{pmatrix}$$

โดยการใช้ตัวดำเนินการแถว

Chapter 3: Application

3.4) ปัญหาค่าเฉพาะ (Eigenvalue problems)

อีกหนึ่งการประยุกต์ใช้ของพีชคณิตเชิงเส้นคือการหาค่าเฉพาะของตัวดำเนินการซึ่งเป็นปัญหาที่เราพบได้บ่อยในฟิสิกส์ สำหรับตัวดำเนินการหนึ่งๆมักจะมีเวกเตอร์และค่าที่บอกถึงลักษณะเฉพาะของการแปลงนั้นๆเสมอ พิจารณาตัวอย่างของตัวดำเนินการสะท้อนรอบแกน x = y ซึ่งเขียนแทนด้วยแมทริกซ์ตัวแทนดังนี้

$$\hat{T} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

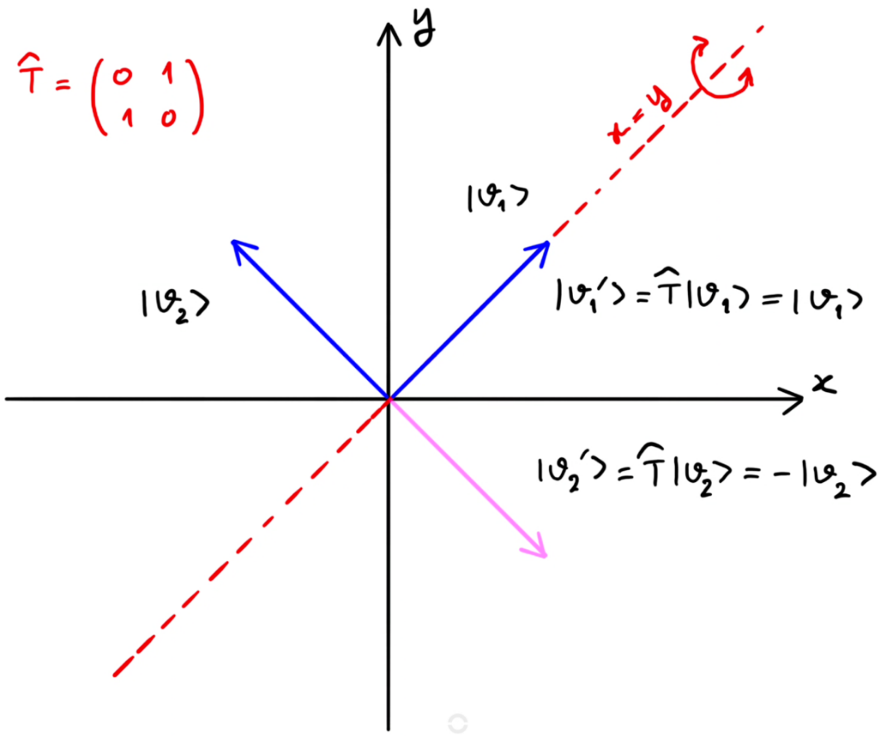
เราพบว่ามีเวกเตอร์สองตัวที่ไม่เปลี่ยนทิศทางภายใต้การแปลงนั่นคือ

$$|v_1\rangle = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rightarrow |v_1'\rangle = \hat{T}|v_1\rangle = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1 \cdot |v_1\rangle$$

และ

$$|v_2\rangle = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \rightarrow |v_2'\rangle = \hat{T}|v_2\rangle = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} = -1 \cdot |v_2\rangle$$

(สังเกตว่าเรามองการคูณด้วยตัวเลข -1 เป็นเวกเตอร์ในทิศทางเดิม แม้ว่าในภาพ 2 มิติมันคือการกลับทิศทางก็ตาม)



โดยเราสามารถเขียนความจริงที่ว่าเวกเตอร์สองตัวนี้ไม่เปลี่ยนทิศทางภายใต้การแปลง เป็นสมการลักษณะเฉพาะ (Eigen equation)

$$\hat{T}|v\rangle = \lambda|v\rangle$$

เวกเตอร์ที่สอดคล้องกับสมการนี้เราเรียกว่า**เวกเตอร์ลักษณะเฉพาะ (Eigenvectors)** และตัวเลขสเกล่าร์ λ ที่คู่กับเวกเตอร์ลักษณะเฉพาะนี้เราเรียกว่า**ค่าลักษณะเฉพาะ (Eigenvalues)** สังเกตว่าถ้าเราให้เวกเตอร์เป็นเวกเตอร์ศูนย์ $|v\rangle = |0\rangle$ เราจะพบว่าเวกเตอร์ศูนย์เป็นคำตอบของทุกตัวดำเนินการ ดังนั้นเราจะไม่นับเวกเตอร์ศูนย์เป็นคำตอบของสมการลักษณะเฉพาะ

เพื่อที่จะหาเวกเตอร์ลักษณะเฉพาะในกรณีทั่วไป เราจะเริ่มพิจารณาจากสมการ

$$\hat{T}|v\rangle = \lambda|v\rangle$$
$$\hat{T}|v\rangle - \lambda\hat{\mathbb{I}}|v\rangle = |0\rangle$$
$$\left(\hat{T} - \lambda\hat{\mathbb{I}}\right)|v\rangle = |0\rangle$$

สมมติว่าถ้าเราสามารถหาอินเวอร์สของ $\left(\hat{T} - \lambda\hat{\mathbb{I}}\right)$ ได้ เราจะพบว่าเมื่อคูณทั้งสองข้างของสมการด้วยอินเวอร์สเราจะได้คำตอบเป็น

$$\left(\hat{T} - \lambda\hat{\mathbb{I}}\right)^{-1}\left(\hat{T} - \lambda\hat{\mathbb{I}}\right)|v\rangle = \left(\hat{T} - \lambda\hat{\mathbb{I}}\right)^{-1}|0\rangle = |0\rangle$$
$$|v\rangle = |0\rangle$$

ซึ่งจะเป็นคำตอบที่เราไม่ต้องการ ดังนั้นวิธีการหาเวกเตอร์ลักษณะเฉพาะคือเราต้องใช้เงื่อนที่ว่า $\left(\hat{T} - \lambda\hat{\mathbb{I}}\right)$ หาอินเวอร์สไม่ได้ นั่นคือ $\left(\hat{T} - \lambda\hat{\mathbb{I}}\right)$ เป็น singular matrix

$$\det\left(\hat{T} - \lambda\hat{\mathbb{I}}\right) = 0$$

ซึ่งเงื่อนไขนี้เราสามารถหาค่าลักษณะเฉพาะจากการหาดีเทอมิแนนท์ แต่อย่างไรก็ดีการหาดีเทอมิแนนท์โดยการใช้แมทริกซ์ตกแต่งและตัวดำเนินการแถวช่วยเราไม่ได้มากนักในกรณีที่สมาชิกของแมทริกซ์เป็นตัวแปร

ดังนั้นเราจะใช้**สูตรของไลบ์นิซ (Leibniz formula for determinant)** แทน

$$\det \hat{A} = \epsilon_{i_1 i_2 \cdots i_n} a_{i_1 1} a_{i_2 2} \cdots a_{i_n n}$$

โดยที่ $\epsilon_{i_1 i_2 \cdots i_n}$ คือ**สัญลักษณ์เลวี-ชีวิต้า (Levi-Civita symbol)**

$$\epsilon_{i_1 i_2 \cdots i_n} = \begin{cases} +1, & even\ permutation \\ -1 & odd\ permutation \\ 0 & repeat\ index \end{cases}$$

โดยมีค่าเป็น $+1$ ถ้าดัชนี $i_1 i_2 \cdots i_n$ เป็นการสลับที่จาก $12 \cdots n$ เป็นจำนวนคู่ครั้ง

และมีค่าเป็น -1 ถ้าดัชนี $i_1 i_2 \cdots i_n$ เป็นการสลับที่จาก $12 \cdots n$ เป็นจำนวนคี่ครั้ง

และมีค่าเป็น 0 ถ้าดัชนี $i_1 i_2 \cdots i_n$ มีการซ้ำกัน

ตัวอย่าง จงหาค่าของ $\epsilon_{11}, \epsilon_{12}, \epsilon_{21}, \epsilon_{22}$ พร้อมทั้งแสดงว่า

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \epsilon_{ij} a_{1i} a_{2j}$$

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \epsilon_{11} a_{11} a_{21} + \epsilon_{12} a_{11} a_{22} + \epsilon_{21} a_{12} a_{21} + \epsilon_{22} a_{12} a_{22} = a_{11} a_{21} - a_{12} a_{21}$$

ตัวอย่าง จงหาค่าของ $\epsilon_{123}, \epsilon_{312}, \epsilon_{231}, \epsilon_{213}, \epsilon_{321}, \epsilon_{132}$ พร้อมทั้งแสดงว่า

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \epsilon_{ijk} a_{1i} a_{2j} a_{3k}$$

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \epsilon_{123} a_{11} a_{22} a_{33} + \epsilon_{321} a_{13} a_{22} a_{31} + \epsilon_{231} a_{12} a_{23} a_{31}$$

$$+\epsilon_{213} a_{12} a_{21} a_{33} + \epsilon_{321} a_{13} a_{22} a_{31} + \epsilon_{132} a_{11} a_{23} a_{32}$$

จากเงื่อนไขที่ว่า $(\hat{T}-\lambda\hat{\mathbb{I}})$ เป็น singular matrix

$$\det(\hat{T}-\lambda\hat{\mathbb{I}})=0$$

สูตรการหาดีเทอมิแนนท์บอกเราว่าถ้า $\hat{T}$ เป็นแมทริกซ์ขนาด $n\times n$ สมการนี้จะเป็นพหุนามดีกรี n (polynomial of degree n^{th}) นั่นคือ

$$P_n(\lambda)=A_n\lambda^n+A_{n-1}\lambda^{n-1}+\cdots+A_1\lambda+A_0=0$$

ซึ่งการแก้สมการนี้จะให้คำตอบทั้งหมด n ราก และแต่ละรากจะสอดคล้องกับเวกเตอร์ลักษณะเฉพาะของตัวเอง โดยการแก้หาเวกเตอร์ลักษณะเฉพาะทำได้โดยการแทนค่าลักษณะเฉพาะกลับเข้าไปในสมการลักษณะเฉพาะแล้วหาค่าเวกเตอร์ที่สอดคล้องกับสมการนั้น

เนื่องจากว่าเวกเตอร์ลักษณะเฉพาะมีขนาดเท่าใดก็ได้ โดยปรกติแล้วเรามักจะเลือกเขียนคำตอบของเวกเตอร์ลักษณะเฉพาะเป็นเวกเตอร์ที่มีขนาดเท่ากับหนึ่ง (normalized vector) เสมอ

ตัวอย่าง จงหาค่าลักษณะเฉพาะของแมทริกซ์ของเพาลี

$$\hat{\sigma}_1=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix},\qquad \hat{\sigma}_2=\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix},\qquad \hat{\sigma}_3=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$$

พร้อมทั้งหาเวกเตอร์ลักษณะเฉพาะ

ตัวอย่าง จงหาค่าลักษณะเฉพาะของ

$$\hat{A} = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$$

พร้อมทั้งหาเวกเตอร์ลักษณะเฉพาะ

ตัวอย่าง จงหาค่าลักษณะเฉพาะของ

$$\hat{T} = \begin{pmatrix} 7 & 0 & -3 \\ -9 & -2 & 3 \\ 18 & 0 & -8 \end{pmatrix}$$

พร้อมทั้งหาเวกเตอร์ลักษณะเฉพาะ

สมบัติของค่าลักษณะเฉพาะและเวกเตอร์ลักษณะเฉพาะที่มีประโยชน์กับฟิสิกส์มีดังต่อไปนี้

1) สำหรับตัวดำเนินการหนึ่ง ถ้าค่าลักษณะเฉพาะแตกต่างกัน เวกเตอร์ลักษณะเฉพาะจะเป็นอิสระเชิงเส้นกัน
2) ถ้าตัวดำเนินการเป็นเฮอมิเทียน ค่าลักษณะเฉพาะจะเป็นจำนวนจริงทั้งหมด

เราจะละเว้นการพิสูจน์ในระดับนี้ แต่จะแสดงตัวอย่างที่เห็นได้บ่อยในฟิสิกส์แทน

ตัวอย่าง จงแสดงว่าแมทริกซ์เพาลีมีเวกเตอร์ลักษณะเฉพาะเป็นอิสระเชิงเส้นกัน และค่าลักษณะเฉพาะเป็นจำนวนจริงทั้งหมด

$$\hat{\sigma}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \hat{\sigma}_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \hat{\sigma}_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

ตัวอย่าง จงแสดงว่าแมทริกซ์ 3×3 ที่เป็นแมทริกซ์ตัวแทนของตัวดำเนินการโมเมนตัมเชิงมุมที่มี $l = 1$ มีเวกเตอร์ลักษณะเฉพาะเป็นอิสระเชิงเส้นกัน และค่าลักษณะเฉพาะเป็นจำนวนจริงทั้งหมด

$$\hat{L}_x = \frac{\hbar}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \hat{L}_y = \frac{\hbar}{\sqrt{2}}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \qquad \hat{L}_z = \hbar\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

3.5) การทำให้แมทริกซ์เป็นแนวทแยง (Diagonalization)

พิจารณาสมการลักษณะเฉพาะของตัวดำเนินการ $\hat{T}$ ขนาด 3×3 ซึ่งมีเวกเตอร์ลักษณะเฉพาะทั้งหมดสามตัวดังนี้

$$\hat{T}|u\rangle = \lambda_u|u\rangle \rightarrow \begin{pmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{22} \\ T_{31} & T_{32} & T_{33} \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \lambda_u\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$$

$$\hat{T}|v\rangle = \lambda_v|v\rangle \rightarrow \begin{pmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{22} \\ T_{31} & T_{32} & T_{33} \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \lambda_v\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

$$\hat{T}|w\rangle = \lambda_w|w\rangle \rightarrow \begin{pmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{22} \\ T_{31} & T_{32} & T_{33} \end{pmatrix}\begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \lambda_w\begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix}$$

เราสามารถเขียนสมการทั้งสามนี้ในรูปของการคูณกันของแมทริกซ์ดังนี้

$$\begin{pmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{22} \\ T_{31} & T_{32} & T_{33} \end{pmatrix}\begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix} = \begin{pmatrix} \lambda_u u_1 & \lambda_v v_1 & \lambda_w w_1 \\ \lambda_u u_2 & \lambda_v v_2 & \lambda_w w_2 \\ \lambda_u u_3 & \lambda_v v_3 & \lambda_w w_3 \end{pmatrix}$$

$$\begin{pmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{22} \\ T_{31} & T_{32} & T_{33} \end{pmatrix}\begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix} = \begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix}\begin{pmatrix} \lambda_u & 0 & 0 \\ 0 & \lambda_v & 0 \\ 0 & 0 & \lambda_w \end{pmatrix}$$

$$\hat{T}\hat{U} = \hat{U}\hat{D}$$

ถ้าเราคูณด้วยอินเวอร์สแมทริกซ์เราจะได้

$$\hat{U}^{-1}\hat{T}\hat{U} = \hat{U}^{-1}\hat{U}\hat{D} = \hat{D}$$

นั่นคือเราพบว่าเราสามารถทำให้ตัวดำเนินการ กลายไปเป็นแมทริกซ์แนวทแยงที่มีค่าลักษณะเฉพาะตามแนวทแยงโดยการนำแมทริกซ์และอินเวอร์สของมัน

$$\boxed{\hat{U} = \begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix}}$$

ซึ่งหาได้จากเวกเตอร์ลักษณะเฉพาะไปคูณ

$$\boxed{\begin{pmatrix} \lambda_u & 0 & 0 \\ 0 & \lambda_v & 0 \\ 0 & 0 & \lambda_w \end{pmatrix} = \hat{D} = \hat{U}^{-1}\hat{T}\hat{U}}$$

กระบวนการนี้เรียกว่าการทำให้แมทริกซ์เป็นแนวทะแยง (diagonalization) ซึ่งเราสามารถประยุกต์ใช้ในสมการทางฟิสิกส์ได้หลากหลาย

ข้อสังเกตของสมการ $\hat{D} = \hat{U}^{-1}\hat{T}\hat{U}$ คือเราจะได้ความสัมพันธ์ของดีเทอมีแนนท์ (determinant) และผลรวมแนวทแยง (trace) ของ $\hat{T}$ ในรูปของค่าลักษณะเฉพาะ

$$\det(\hat{D}) = \lambda_u\lambda_v\lambda_w = \prod_i \lambda_i = \det(\hat{U}^{-1}\hat{T}\hat{U}) = \frac{1}{\det(\hat{U})}\det(\hat{T})\det(\hat{U}) = \det(\hat{T})$$

$$\boxed{\det(\hat{T}) = \prod_i \lambda_i}$$

$$Tr(\hat{D}) = \lambda_u + \lambda_v + \lambda_w = \sum_i \lambda_i = Tr(\hat{U}^{-1}\hat{T}\hat{U}) = Tr(\hat{T}\hat{U}\hat{U}^{-1}) = Tr(\hat{T})$$

$$\boxed{Tr(\hat{T}) = \sum_i \lambda_i}$$

ตัวอย่าง พิจารณาสมการของ coupled harmonic oscillators (q_1, q_2)

$$\ddot{q}_1 + q_1 + 2q_2 = 0$$
$$\ddot{q}_2 + 2q_1 + q_2 = 0$$

ซึ่งเขียนเป็นสมการเวกเตอร์ได้เป็น

$$\frac{d^2}{dt^2}\begin{pmatrix} q_1 \\ q_2 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} q_1 \\ q_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

จงหาผลรวมเชิงเส้นของ (q_1, q_2) ที่มีการสั่นอย่างอิสระ

ตัวอย่าง จงเปลี่ยนแมทริกซ์ต่อไปนี้ให้เป็นแมทริกซ์แนวทแยง

1) $\hat{\sigma}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

2) $\hat{\sigma}_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$

3) $\hat{L}_x = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$

4) $\hat{L}_y = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}$

5) $\hat{A} = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$

แบบฝึกหัด 3

1) ระบบสมการ

1.1) จงแก้ระบบสมการต่อไปนี้โดยใช้ตัวดำเนินการแถว

$$y - z = 3$$
$$-2x + 4y - z = 1$$
$$-2x + 5y - 4z = -2$$

1.2) จงแก้ระบบสมการต่อไปนี้โดยใช้ตัวดำเนินการแถว

$$2x + y + z = 1$$
$$6x + 2y + z = -1$$
$$-2x + 2y + z = 7$$

1.3) จงแก้ระบบสมการต่อไปนี้โดยใช้ตัวดำเนินการแถว

$$-3x - 3y + 9z = 12$$
$$2x + 2y - 4z = -2$$
$$-2y - 4z = -8$$

2) การหาอินเวอร์สของแมทริกซ์

2.1) จงหาอินเวอร์สของแมทริกซ์ด้านล่างโดยใช้ตัวดำเนินการแถว

$$\hat{A} = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 1 & 1 \\ -1 & 1 & 0 \end{pmatrix}$$

2.2) จงหาอินเวอร์สของแมทริกซ์ด้านล่างโดยใช้ตัวดำเนินการแถว

$$\hat{A} = \begin{pmatrix} 0 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 0 \end{pmatrix}$$

2.3) จงหาอินเวอร์สของแมทริกซ์ด้านล่างโดยใช้ตัวดำเนินการแถว

$$\hat{A} = \begin{pmatrix} 7 & 2 & 1 \\ 0 & 3 & -1 \\ -3 & 4 & -2 \end{pmatrix}$$

3) การหาดีเทอมิแนนท์

3.1) จงหาดีเทอมิแนนท์ของแมทริกซ์ด้านล่างโดยใช้ตัวดำเนินการแถว

$$\hat{A} = \begin{pmatrix} 7 & 2 & 1 \\ 0 & 3 & -1 \\ -3 & 4 & -2 \end{pmatrix}$$

3.2) จงหาดีเทอมิแนนท์ของแมทริกซ์ด้านล่างโดยใช้ตัวดำเนินการแถว

$$\hat{A} = \begin{pmatrix} 1 & -2 & 3 \\ 2 & 0 & 3 \\ 1 & 5 & 4 \end{pmatrix}$$

3.3) จงหาดีเทอมิแนนท์ของแมทริกซ์ด้านล่างโดยใช้ตัวดำเนินการแถว

$$\hat{A} = \begin{pmatrix} 2 & -3 & 1 \\ 2 & 0 & -1 \\ 1 & 4 & 5 \end{pmatrix}$$

4) การหาค่าลักษณะเฉพาะและเวกเตอร์ลักษณะเฉพาะ

4.1) จงหาค่าลักษณะเฉพาะและเวกเตอร์ลักษณะเฉพาะของตัวดำเนินการ

$$\hat{A} = \begin{pmatrix} 0 & 1 \\ -2 & -3 \end{pmatrix}$$

4.2) จงหาค่าลักษณะเฉพาะและเวกเตอร์ลักษณะเฉพาะของตัวดำเนินการ

$$\hat{A} = \begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix}$$

4.3) จงหาค่าลักษณะเฉพาะและเวกเตอร์ลักษณะเฉพาะของตัวดำเนินการ

$$\hat{A} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

5) การทำให้เป็นแนวทแยง

5.1) จงทำให้ตัวดำเนินการต่อไปนี้เป็นแมทริกซ์แนวทะแยง

$$\hat{A} = \begin{pmatrix} 0 & 1 \\ -2 & -3 \end{pmatrix}$$

5.2) จงทำให้ตัวดำเนินการต่อไปนี้เป็นแมทริกซ์แนวทะแยง

$$\hat{A} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

5.3) จงทำให้ตัวดำเนินการต่อไปนี้เป็นแมทริกซ์แนวทะแยง

$$\hat{A} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & 2 \end{pmatrix}$$

Chapter 4: Basic Statistics

4.1) Probability and Random Variables (ความน่าจะเป็นและตัวแปรสุ่ม)

แนวคิดที่สำคัญของคณิตศาสตร์ที่ใช้ในฟิสิกส์คือ**ความไม่แน่นอน (uncertainty)** ซึ่งมีการแสดงรูปแบบได้หลากหลายทาง ตัวอย่างเช่น เรามักจะพบกับสถานการณ์ที่ผลของการวัด (measurement) เปลี่ยนแปลงไปเมื่อทำการทดลองซ้ำ โดยพฤติกรรมเหล่านี้อาจจะเกิดจากความผิดพลาดของเครื่องมือวัด หรืออาจจะเป็นผลมาจากกฎเกณฑ์ของธรรมชาติเอง (ปรากฏการณ์ในระดับควอนตัม) ที่ทำให้ไม่สามารถทำนายระบบได้อย่างแม่นยำ ไม่ว่าจะมาจากสาเหตุใด ลักษณะเฉพาะของระบบเมื่อเราไม่สามารถทำนายผลลัพท์ได้อย่างแม่นยำแบบนี้ เราจะเรียกว่ามี**สมบัติสุ่ม (random)**
โดยการวัดผลของการสุ่มในระบบนั้นเราจะเรียกว่า**ความน่าจะเป็น (probability)** โดยเราจะเริ่มนิยามความน่าจะเป็นผ่านทางทฤษฎีเซต
พิจารณาเซต S ที่เราจะเรียกว่า**ปริภูมิตัวอย่าง (sample space)** ซึ่งมีจำนวนสมาชิกที่แน่นอน เราจะกำหนดว่าทุกเซตย่อย (subset) A ของ S ใดๆจะมีค่าจำนวนจริง $P(A)$ คู่อยู่กับมันเสมอและมีสมบัติต่อไปนี้

1. ค่าของ $P(A)$ สำหรับเซตย่อย A ใดๆ มากกว่าศูนย์เสมอ $P(A) > 0$
2. ถ้าเซตย่อย A และเซตย่อย B ไม่มีสมาชิกซ้ำกันเลยเราจะมี $P(A \cup B) = P(A) + P(B)$
3. ค่า $P(S) = 1$

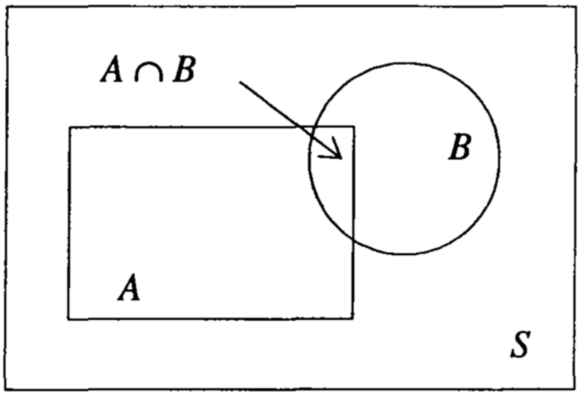
ซึ่งเราสามารถใช้สมบัติทั้งสามเพื่อแสดงสมบัติพื้นฐานที่เหลือได้ดังต่อไปนี้

1. $P(A') = 1 - P(A)$ โดย A' คือคอมพลีเมนต์ของเซต A
2. $P(A \cup A') = 1$
3. $P(\emptyset) = 0$
4. ถ้า $A \subset B$ แล้ว $P(A) \leq P(B)$
5. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

ส่วนตัวแปรซึ่งมีค่าสำหรับแต่ละสมาชิกของปริภูมิตัวอย่างเราจะเรียกว่าตัวแปรสุ่ม (random variable) โดยตัวแปรเหล่านี้อาจจะมีหลายมิติซึ่งอาจเขียนแทนด้วยเวกเตอร์ได้

สำหรับเซตย่อย A และ B ของปริภูมิตัวอย่าง S ถ้า $P(B) \neq 0$ แล้วเราจะนิยามความน่าจะเป็นแบบมีเงื่อนไข (conditional probability) $P(A|B)$ จาก

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$



สังเกตว่าเราอาจจะมอง $P(A)$ ได้จาก $P(A) \ = \ P(A|S)$ และเซตย่อย A และ B จะเป็นอิสระต่อกันเมื่อ

$$P(A \cap B) = P(A)P(B)$$

และเพราะว่า $P(A \cup B) = P(A|B)P(B)$ และ $P(B \cup A) = P(B|A)P(A)$
เราจะได้

$$P(A|B) = \frac{P(B|A)\ P(A)}{P(B)}$$

สมการนี้เราจะเรียกว่า**ทฤษฎีของเบยส์ (Bayes’ theorem)**

พิจารณาปริภูมิตัวอย่าง S ที่มีเซตย่อย A_i ซึ่งไม่มีสมาชิกซ้ำกันเลย (disjoint subsets) นั่นคือ $S = \bigcup_i A_i$ และ $A_i \cap A_j = \emptyset$ ถ้าเราพิจารณาเซตย่อย B เราจะพบว่า $B \cap A_i$ ไม่มีสมาชิกซ้ำกันเช่นกัน ดังนั้นเราจะได้

$$\boxed{P(B) = P\left(\bigcup_i B \cap A_i\right) = \sum_i P(B \cap A_i) = \sum_i P(B|A_i)P(A_i)}$$

ซึ่งเราจะเรียกสมการนี้ว่ากฎของความน่าจะเป็นรวม (Law of total probability) ซึ่งจะมีประโยชน์ตอนเราพิจารณาการแตกปริภูมิตัวอย่างเป็นเซตย่อยที่ไม่มีสมาชิกซ้ำกัน ตัวอย่างเช่น ทฤษฎีของเบยส์จะเขียนได้เป็น

$$\boxed{P(A|B) = \frac{P(B|A)\, P(A)}{\sum_i P(B|A_i)P(A_i)}}$$

โดยที่เซตย่อย A เป็นเซตใดๆ

ตัวอย่าง จงสร้างปริภูมิตัวอย่างของการโยนเหรียญ และนิยามความน่าจะเป็นให้อย่างเหมาะสม

ตัวอย่าง จงสร้างปริภูมิตัวอย่างของการทอยลูกเต๋าพร้อมทั้งนิยามความน่าจะเป็นอย่างเหมาะสม และคำนวณหาความน่าจะเป็นที่ลูกเต๋ามีหน้าเป็นเลขคู่

ตัวอย่าง จงสร้างปริภูมิตัวอย่างของการโยนเหรียญสองเหรียญ (ไม่คำนึงถึงอันดับ) และนิยามความน่าจะเป็นให้อย่างเหมาะสม จากนั้นให้คำนวณความน่าจะเป็นที่เหรียญหงายหน้าไม่เหมือนกัน

ตัวอย่าง จงสร้างปริภูมิตัวอย่างของการทอยลูกเต๋าพร้อมกันสองลูก (ไม่คำนึงถึงอันดับ) และนิยามความน่าจะเป็นให้อย่างเหมาะสม จากนั้นให้คำนวณหาความน่าจะเป็นที่ลูกเต๋ามีผลรวมเป็น 3

ตัวอย่าง โรคชนิดทางพันธุกรรมหนึ่งซึ่งเราทราบว่ามีโอกาสเป็น 0.1% ในประชากร นั่นคือเราจะมีความน่าจะเป็นก่อน (prior probability) เป็น

$$P\left(\text{เป็นโรค}\right) = 0.001$$

$$P\left(\text{ไม่เป็นโรค}\right) = 0.999$$

ถ้ามีชุดทดสอบชนิดหนึ่งซึ่งให้ผลเป็นบวก (ผลแสดงว่าเป็นโรค) 98% ถ้าผู้ทดสอบเป็นโรค นั่นคือ

$$P\left(+|\text{เป็นโรค}\right) = 0.98$$

$$P\left(-|\text{เป็นโรค}\right) = 0.02$$

และกำหนดให้มีโอกาส 3% ที่จะได้ผลเป็นบวกสำหรับผู้ทดสอบที่ไม่เป็นโรค นั่นคือ

$$P\left(+|\text{ไม่เป็นโรค}\right) = 0.03$$

$$P\left(-|\text{ไม่เป็นโรค}\right) = 0.97$$

จงหาความน่าจะเป็นภายหลัง (posterior probability) ที่บอกว่าถ้าผลทดสอบรายงานผลเป็นบวกแล้วจะมีความน่าจะเป็นเท่าใดที่จะเป็นโรคนั้นจริง นั่นคือจงหาว่า

$$P\left(\text{เป็นโรค}|+\right) =?$$

4.2) การตีความของความน่าจะเป็น

ถึงแม้ว่าฟังก์ชั่นใดๆบนปริภูมิตัวอย่างที่สามารถใช้แทนความน่าจะเป็นได้ (ตามนิยามด้านบน) เรายังต้องหาวิธีตีความสมาชิกของปริภูมิตัวอย่างและวิธีในการใส่ค่าความน่าจะเป็นและตีความความน่าจะเป็น มีสองวิธีในการตีความของความน่าจะเป็นที่เราจะพบได้บ่อยในการวิเคราะห์ทางสถิติและวิเคราะห์ข้อมูล วิธีแรกคือ**ความถี่สัมพัทธ์ (relative frequency)** ซึ่งนิยมใช้ในการคำนวณความผิดพลาด (statistical errors) ของวัด อีกวิธีหนึ่งคือ**ความน่าจะเป็นเชิงอัตวิสัย (subjective probability)** ซึ่งนิยมใช้ในการหาความผิดพลาดเชิงระบบ (systematic errors)

มุมมองความถี่สัมพัทธ์ (Relative Frequency)

ในมุมมองนี้เราจะสมมติให้ปริภูมิตัวอย่าง S แทนที่ผลลัพธ์ที่เป็นไปได้ทั้งหมดของการวัด ซึ่งการวัดนี้มีลักษณะที่ทำซ้ำได้ (repeatable) เซตย่อย A ของ S ที่สอดคล้องกับผลลัพธ์บางประเภทเราจะเรียกเซตย่อยนี้ว่าเหตุการณ์ (event) เราจะเริ่มนิยามความน่าจะเป็นจากการให้ค่าความน่าจะเป็นกับสมาชิกแต่ละตัวโดยการพิจารณาเซตย่อยที่มีสมาชิกเพียงตัวเดียว โดยเราจะให้ความน่าจะเป็นนี้มีค่าเท่ากับสัดส่วนของผลลัพท์ที่อยู่ในเซตย่อยนี้หารด้วยจำนวนการวัดทั้งหมดและวัดเป็นจำนวนอนันต์ครั้งนั่นคือ

$$P(A) = \lim_{n\to\infty} \frac{\text{จำนวนครั้งที่ผลลัพท์เป็น } A \text{ จากการวัด } n \text{ ครั้ง}}{n}$$

ส่วนความน่าจะเป็นสำหรับเซตย่อยใดๆเราสามารถนำค่าความน่าจะเป็นของเซตย่อยที่มีสมาชิกตัวเดียวมาบวกกันได้

ในกรณีของความน่าจะเป็นแบบมีเงื่อนไขเรานิยามมันได้จาก

$$P(A|B) = \lim_{n\to\infty} \frac{\text{จำนวนครั้งที่ผลลัพท์เป็น } A \text{ และ } B \text{ จากการวัด } n \text{ ครั้ง}}{\text{จำนวนครั้งที่ผลลัพธ์เป็น } B \text{ จากการวัด } n \text{ ครั้ง (ไม่สนใจว่าเป็น } A \text{ หรือไม่)}}$$

เราจะเห็นได้ชัดว่าวิธีการตีความนี้ไม่สามารถให้ความสมบูรณ์แบบได้จากการทดลอง (ต้องทำการทดลองอนันต์ครั้ง) และทำให้เราต้องทำการประมาณเพื่อหาความน่าจะเป็นจากข้อมูลที่มีจำกัด

และเราทำได้เพียงทดสอบว่าแบบจำลอง (model) หรือทฤษฎีที่เราสร้างขึ้นให้ความน่าจะเป็นที่สอดคล้องกับการวัดหรือไม่
การตีความแบบนี้นิยมใช้กับการศึกษากฎทางฟิสิกส์ ซึ่งจะมีลักษณะเดียวกันไม่ว่าจะทำการทดลองซ้ำกี่ครั้ง ตัวอย่างเช่นการชนกันของอนุภาคในเครื่องเร่งอนุภาคซึ่งเราสามารถทำการชนของอนุภาคชนิดเดิมซ้ำๆได้ แต่อย่างไรก็ตามแนวคิดแบบนี้ไม่เหมาะกับเหตุการณ์ที่มีเพียงครั้งเดียวและไม่สามารถทำซ้ำได้ เช่นบิ๊กแบงตอนกำเนิดเอกภพเป็นต้น

<u>มุมมองอัตวิสัยหรือมุมมองเบย์เซียน (Subjective/Bayesian Probability)</u>

อีกมุมมองหนึ่งคือมุมมองอัตวิสัย (subjective) ซึ่งเป็นมุมมองที่ขึ้นกับความเชื่อของผู้ทำการวิเคราะห์ ในมุมมองนี้สมาชิกของปริภูมิตัวอย่างจะสอดคล้องกับสมมติฐาน (hypotheses) หรือข้อเสนอ (proposition) ซึ่งเป็นประพจน์ที่มีค่าจริงหรือเท็จ และเราจะตีความความน่าจะเป็นจาก

$$P(A) = \text{ระดับความเชื่อมั่นว่าสมมติฐาน } A \text{ เป็นจริง}$$

โดยปริภูมิตัวอย่างต้องสร้างจากสมมติฐานที่ไม่ซ้อนทับกันนั่นคือมีเพียงสมมติฐานเดียวในปริภูมิตัวอย่างเท่านั้นที่เป็นจริง
ข้อความที่บอกว่าการวัดจะได้ผลลัพธ์เป็นจำนวนครั้งที่แน่นอนสามารถถูกพิจารณาให้เป็นสมมติฐานได้เช่นกัน (ข้อความนี้อาจจะถูกหรือผิด) นั่นคือมุมมองนี้ครอบคลุมถึงมุมมองความถี่สัมพัทธ์ด้วย
มากไปกว่านั้นมุมมองอัตวิสัยสามารถนำไปเกี่ยวข้องกับการหาค่าคงที่ที่ไม่ทราบค่าในธรรมชาติได้ ซึ่งมันจะแสดงถึงความมั่นใจว่าค่าคงที่นั้นจะอยู่ในช่วงของค่าช่วงหนึ่ง ความน่าจะเป็นของค่าคงที่เหล่านี้ไม่สามารถมองจากมุมมองความถี่สัมพัทธ์ได้ ตัวอย่างเช่นถ้าเราวัดมวลของอิเล็กตรอนเราจะไม่สามารถให้ความน่าจะเป็นของเหตุการณ์ที่มวลอิเล็กตรอนเท่ากับ $1\,\frac{GeV}{c^2}$ ได้เป็นต้น
การใช้มุมมองอัตวิสัยนั้นมีความใกล้เคียงกับทฤษฎีของเบยส์มาก และนี่เป็นพื้นฐานของสถิติเบย์เซียน เซตย่อย A ซึ่งปรากฏในสมการ

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

อาจถูกพิจารณาให้เป็นสมมติฐานที่ทฤษฎีอันหนึ่งเป็นจริง และเซตย่อย B เป็นสมมติฐานที่บอกว่าการทดลองจะให้ผลการทดลองที่แน่นอนค่าหนึ่ง (มาจากข้อมูล) สมการด้านบนจึงมักจะเขียนในรูป

$$\boxed{P(\text{ทฤษฎี}|\text{ข้อมูล}) \propto P(\text{ข้อมูล}|\text{ทฤษฎี})\, P(\text{ทฤษฎี})}$$

ในที่นี้เราจะเรียก $P(\text{ทฤษฎี})$ ซึ่งเป็นความน่าจะเป็นก่อน (Prior probability) ที่ทฤษฎีเป็นจริง และ $P(\text{ข้อมูล}|\text{ทฤษฎี})$ ซึ่งเราเรียกว่าความเป็นไปได้ (likelihood) ที่บอกเราว่าถ้าสมมติให้ทฤษฎีจริงแล้วความน่าจะเป็นที่ได้จะข้อมูลตามที่กำหนดให้จะมีค่าเท่าไร ส่วน $P(\text{ทฤษฎี}|\text{ข้อมูล})$ เราเรียกว่าความน่าจะเป็นหลัง (Posterior probability) บอกเราว่าถ้าให้ข้อมูลตามที่กำหนดให้ความน่าจะเป็นที่ทฤษฎีจะจริงมีค่าเท่าไร สังเกตว่าในมุมมองนี้ $P(\text{ทฤษฎี})$ เป็นค่าที่ขึ้นกับผู้ทำการวิเคราะห์และไม่มีกฎตายตัวว่าต้องเลือกอย่างไร แต่ถ้าได้ถูกเลือกแล้วมุมมองนี้บอกเราว่าระดับความเชื่อมั่นของผู้วิเคราะห์เปลี่ยนไปอย่างไรเมื่อพิจารณาจากข้อมูล

ลองพิจารณาตัวอย่างของโรคทางพันธุกรรม สำหรับความน่าจะเป็นที่มีโรคเมื่อทำการทดสอบแล้วเป็นบวกสามารถหาได้จากการศึกษาตัวอย่างประชากรจำนวนมากซึ่งคือการตีความความน่าจะเป็นจากมุมมองความถี่สัมพัทธ์

ในทางกลับกันคนทั่วไปอาจจะสนใจในมุมมองอัตวิสัยมากกว่าเพราะ สำหรับบุคคลทั่วไปไม่สามารถรู้ถึงค่า $P(\text{เป็นโรค})$ ได้ ดังนั้นเราจะมอง $P(\text{เป็นโรค})$ เป็นความน่าจะเป็นก่อน (posterior probaibilty) ที่บอกความมั่นใจของผู้ทำการวิเคราะห์ว่าเค้ามีโอกาสเป็นโรคเท่าใดก่อนเริ่มทำการทดสอบ นี่จึงทำให้มุมมองนี้ขึ้นกับระดับความเชื่อมั่นของคนนั้นๆและเป็นอัตวิสัย ซึ่งทฤษฎีของเบยส์จะบอกเราว่าความน่าจะเป็นที่คนนั้นจะติดโรคเปลี่ยนไปอย่างไรเมื่อได้ทำการทดสอบ

Chapter 4: Basic Statistics

4.3) ฟังก์ชันความหนาแน่นของความน่าจะเป็น

ในหลายครั้งเรามักจะพบกับผลลัพธ์ของการวัดที่เป็นค่าต่อเนื่อง x นั่นคือเราพิจารณาปริภูมิตัวอย่างเป็นค่าต่อเนื่องที่เป็นไปได้จากการวัดนั้นๆ เราอาจจะคิดว่าผลลัพธ์ที่ได้เป็นตัวแปรสุ่มด้วย ในกรณีนี้การแทนค่าน่าจะเป็นจะกลายเป็นการพิจารณาความน่าจะเป็นที่ตัวแปรนี้อยู่ในช่วง x ถึง $x + dx$ และเราจะนิยามความน่าจะเป็นนี้ด้วย**ฟังก์ชันความหนาแน่นของความน่าจะเป็น (probability density function, p.d.f.)** $f(x)$ ตามสมการ

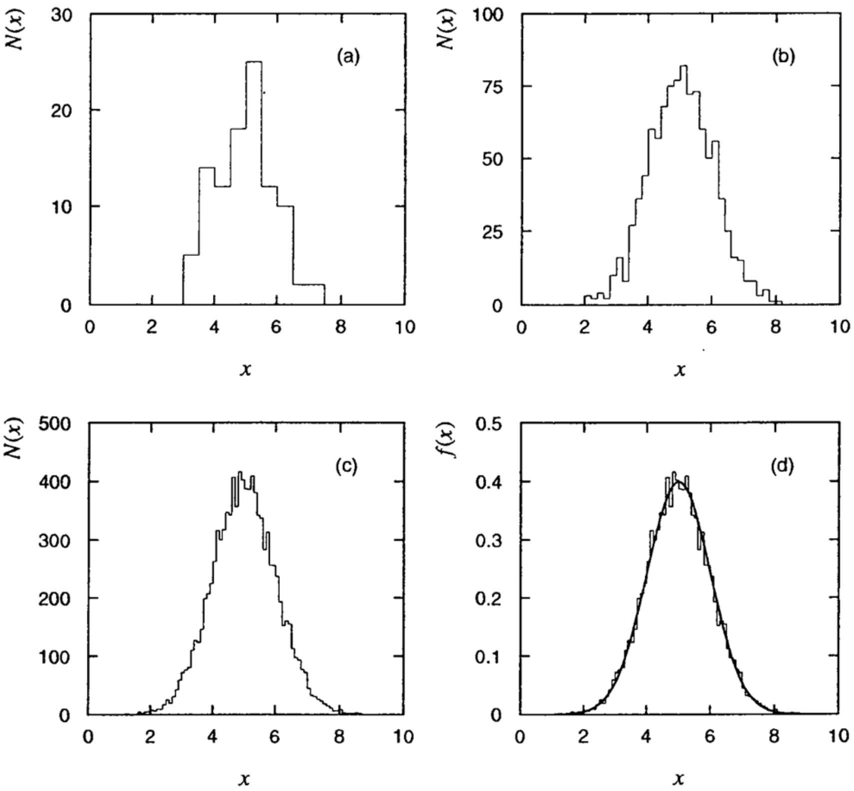
$$P([x, x + \mathrm{d}x]) = f(x)dx$$

ในมุมมองของความถี่สัมพัทธ์ เราตีความว่า $f(x)dx$ เป็นจำนวนครั้งที่เราวัดได้ค่าในช่วง x ถึง $x + dx$ เมื่อเทียบกับจำนวนครั้งในการวัดทั้งหมด เมื่อเราวัดไปเป็นจำนวนอนันต์ครั้ง โดยฟังก์ชั่นความหนาแน่นของความน่าจะเป็นต้องถูก normalize นั่นคือ

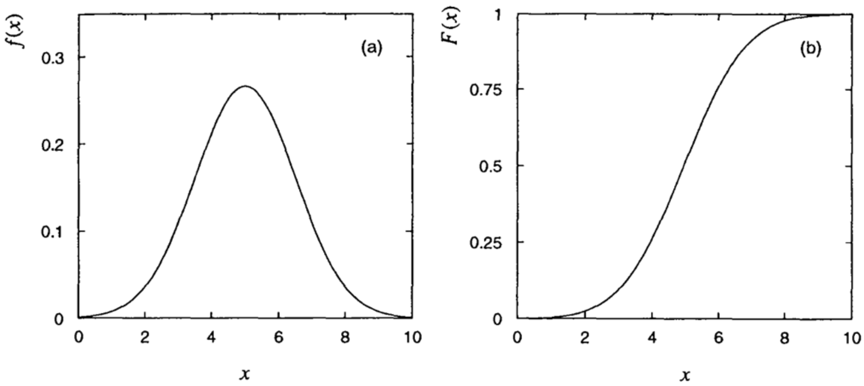
$$\int_S f(x)dx = 1$$

โดยที่ S คือปริภูมิตัวอย่าง (ค่าที่เป็นไปได้ทั้งหมดของ x)

เพื่อที่จะแสดงความสัมพันธ์ของ p.d.f และการวัดจริง $x_1, x_2, \ldots, x_n$ เราจะพิจารณา**ฮิสโตแกรม (histogram)** ของข้อมูลโดยเราแบ่งแกนนอนออกเป็นช่วงๆทั้งหมด m ช่วง โดยแต่ละช่วงเราจะเรียกว่า**แท่ง (bin)** มีระยะห่าง $\Delta x_1, \Delta x_2, \ldots, \Delta x_m$ ซึ่งโดยปรกติเรามักจะแบ่งให้ทุกช่วงเท่ากันทั้งหมด $\Delta x_i = \Delta x$ และความสูงของแท่งในแต่ละช่วง n_i แสดงถึงจำนวนข้อมูลที่อยู่ในช่วง Δx_i นั้น พื้นที่ใต้กราฟของฮิสโตแกรมนี้คือ $\sum_{i=1}^{m} n_i \Delta x_i$ ซึ่งถ้าเราต้องการทำให้พื้นที่ใต้กราฟทั้งหมดนี้ถูก normalize เราต้องหารความสูงด้วย $N\Delta x_i$ ซึ่งจะทำให้พื้นที่ใต้กราฟมีค่าเท่ากับ 1 และกลายเป็น p.d.f. เมื่อการแบ่งแท่งมีระยะห่างเล็กลงมากๆตามรูปด้านล่าง



เมื่อพิจารณาให้ $n \rightarrow \infty, m \rightarrow \infty$ เราจะได้กราฟที่ต่อเนื่องดังภาพ



เราสามารถนิยาม**ฟังก์ชันการแจกแจงความน่าจะเป็นแบบสะสม (cumulative distribution function, c.d.f.)** ได้จาก

$$\boxed{F(x) = \int_{-\infty}^{x} f(x')dx'}$$

ซึ่งเราจะตีความว่า $F(x)$ เป็นความน่าจะเป็นที่ผลลัพธ์จะมีค่าน้อยกว่า x

อีกแนวคิดที่มีประโยชน์มากสำหรับ c.d.f. คือ **ควอนไทล์ที่ p (Quantile of order p, p^{th} quantile)** ซึ่งคือค่า x_p ที่ทำให้ c.d.f มีค่าเท่ากับ p โดยที่ $0 < p \leq 1$ นั่นคือ

$$x_p = F^{-1}(p)$$

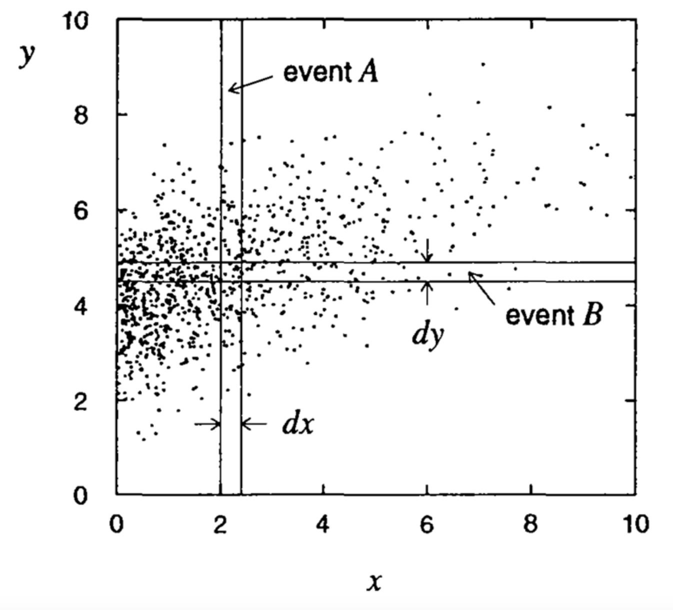
ตัวอย่างที่มักจะใช้แนวคิดของควอนไทล์คือ**ค่ามัธยฐาน (median)**

$$x_{1/2} = F^{-1}\left(\frac{1}{2}\right)$$

ส่วนแนวคิดของ**ฐานนิยม (mode)** สามารถหาได้จากตำแหน่ง x ที่มี p.d.f. สูงสุดซึ่งแสดงถึงตำแหน่งที่มีการซ้ำของข้อมูลเยอะที่สุด

ในกรณีของตัวแปรสุ่มหลายมิติ นั่นคือการวัดหนึ่งครั้งให้ค่าตัวแปร x, y ออกมา ในกรณีนี้เราจะให้ความหมายของเหตุการณ์ A คือเหตุการณ์ที่เราวัดค่า x อยู่ในช่วง x ถึง $x + dx$ และ y มีค่าเป็นอะไรก็ได้ ในขณะที่เหตุการณ์ B คือเหตุการณ์ที่เราวัดค่า y อยู่ในช่วง y ถึง $y + dy$ และ x มีค่าเป็นอะไรก็ได้ เราจะนิยาม **p.d.f. ร่วม (joint p.d.f.)** จาก

$$P(A \cap B) = f(x, y)\, dx\, dy$$



โดยมีเงื่อนไข normalization เป็น

$$\iint_S f(x,y)dxdy = 1$$

สมมติว่าเราทราบ joint p.d.f. แล้วเราอาจจะอยากหา p.d.f. ของ x โดยไม่สนใจว่าค่า y จะเป็นอย่างไร นั่นคือเราอยากหา p.d.f. ของเหตุการณ์ A

$$P(A) = \text{ความน่าจะเป็นที่ } x \text{ อยู่ในช่วง } [x, x+dx] \text{ และ } y \text{ มีค่าเป็นเท่าไรก็ได้}$$
$$P(A) = f_x(x)dx$$

ดังนั้นเราจะได้ **p.d.f. ตามขอบ (marginal p.d.f.)** สำหรับ x เป็น

$$f_x(x) = \int_{-\infty}^{\infty} f(x,y)dy$$

ในทำนองเดียวกัน marginal p.d.f. สำหรับ y เป็น

$$f_y(y) = \int_{-\infty}^{\infty} f(x,y)dx$$

จากความน่าจะเป็นแบบมีเงื่อนไขเราจะได้

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{f(x,y)dxdy}{f_x(x)dx} = \frac{f(x,y)dy}{f_x(x)}$$

ซึ่งจะทำให้เรานิยาม p.d.f แบบมีเงื่อนไข (conditional p.d.f.) ได้เป็น

$$h(y|x) = \frac{f(x,y)}{f_x(x)} = \frac{f(x,y)}{\int f(x,y')dy'}$$

ฟังก์ชันนี้มีความหมายเป็น p.d.f. ของตัวแปร y เพียงตัวแปรเดียวและเรามอง x เป็นค่าคงที่

ในทำนองเดียวกัน p.d.f. แบบมีเงื่อนไขของตัวแปร x โดยที่ y เป็นค่าคงที่จะเขียนได้เป็น

$$g(x|y) = \frac{f(x,y)}{f_y(y)} = \frac{f(x,y)}{\int f(x',y)dx'}$$

และโดยการใช้ทฤษฎีของเบยส์เราจะได้ความสัมพันธ์ระหว่าง p.d.f.แบบมีเงื่อนไขตามสมการ

$$\boxed{g(x|y) = \frac{h(y|x)f_x(x)}{f_y(y)}}$$

และถ้าใช้ความสัมพันธ์ $f(x,y) = h(y|x)f_x(x) = g(x|y)f_y(y)$ เราจะได้ความสัมพันธ์ระหว่าง p.d.f.ตามขอบเป็น

$$\boxed{\begin{aligned} f_x(x) &= \int_{-\infty}^{\infty} g(x|y)f_y(y)dy \\ f_y(y) &= \int_{-\infty}^{\infty} h(y|x)f_x(x)dx \end{aligned}}$$

ในกรณีที่เหตุการณ์ A ซึ่งคือเหตุการณ์ที่ค่า x อยู่ในช่วง x ถึง $x + dx$ และ y มีค่าเป็นอะไรก็ได้ และเหตุการณ์ B ซึ่งคือเหตุการณ์ที่ค่า y อยู่ในช่วง y ถึง $y + dy$ และ x มีค่าเป็นอะไรก็ได้ เป็นอิสระต่อกัน **(A และ B เป็นอิสระต่อกัน)** เราจะได้ว่า $P(A \cap B) = P(A)P(B)$ นั่นคือในรูปของ p.d.f. ร่วมและ p.d.f. ตามขอบจะมีความสัมพันธ์เป็น

$$\boxed{f(x,y) = f_x(x)f_y(y)}$$

ซึ่งเราตีความหมายได้ว่าในกรณีนี้การทราบถึงข้อมูลในตัวแปรหนึ่งไม่เปลี่ยนความน่าจะเป็นของตัวแปรหนึ่ง

ตัวอย่าง พิจารณา p.d.f. ด้านล่าง

$$f(x) = \begin{cases} cx^2, & |x| \leq 1 \\ 0, & \text{กรณีอื่น} \end{cases}$$

1. จงหาค่า c
2. จงหาความน่าจะเป็นที่ค่า x อยู่ในช่วง $[-1, -0.5]$
3. จงหา c.d.f. ของ p.d.f. นี้
4. จงหาควอนไทล์ที่ 0.5

ตัวอย่าง พิจารณา p.d.f. ด้านล่าง

$$f(x) = \begin{cases} 4x^3, & 0 \leq x \leq 1 \\ 0, & \text{กรณีอื่น} \end{cases}$$

1. จงหาความน่าจะเป็นที่ค่า x อยู่ในช่วง $[0,0.5]$
2. จงหา c.d.f. ของ p.d.f. นี้
3. จงหาควอนไทล์ที่ 0.5

ตัวอย่าง พิจารณา p.d.f. ของตัวแปร (x, y) ตามสมการ

$$f(x,y) = \begin{cases} cxy, & 0 \leq x \leq 1, 0 \leq y \leq \sqrt{x} \\ 0, & \text{กรณีอื่น} \end{cases}$$

1. จงหาค่า c
2. จงหา marginal p.d.f. ของตัวแปร x $(f_x(x))$
3. จงหา marginal p.d.f. ของตัวแปร y $(f_y(y))$
4. พิจารณาว่าตัวแปรสุ่ม x และ y เป็นอิสระกันหรือไม่

ตัวอย่าง พิจารณา conditional p.d.f. ของตัวแปร y ตามสมการ

$$h(y|x) = \begin{cases} \frac{1}{2x}, & -x \leq y \leq x \\ 0, & \text{กรณีอื่น} \end{cases}$$

และกำหนดให้ marginal p.d.f. ของตัวแปร x เป็นไปตามสมการ

$$f_x(x) = \begin{cases} 2x, & 0 \leq x \leq 1 \\ 0, & \text{กรณีอื่น} \end{cases}$$

1. จงหา p.d.f ของสองตัวแปร $(f(x, y))$
2. จงหา marginal p.d.f. ของตัวแปร y $(f_y(y))$

Chapter 4: Basic Statistics

4.4) ฟังก์ชันของตัวแปรสุ่ม (Functions of Random Variables)

ฟังก์ชันของตัวแปรสุ่มก็เป็นตัวแปรสุ่มด้วยเช่นกัน สมมติให้ฟังก์ชัน $a(x)$ เป็นฟังก์ชันต่อเนื่องของตัวแปรสุ่ม x โดยมี p.d.f. เป็น $f(x)$ สิ่งที่เราอยากทราบคือ p.d.f. ของตัวแปร a จะมีลักษณะเป็นอย่างไร ถ้าเราใช้เงื่อนไขที่ว่าความน่าจะเป็นที่ x อยู่ในช่วง x ถึง $x + dx$ ตรงกันกับความน่าจะเป็นที่ a อยู่ในช่วง a ถึง $a + da$ ที่สอดคล้องกัน นั่นคือ

$$a = a(x) \text{ และ } a + da = a(x + dx)$$

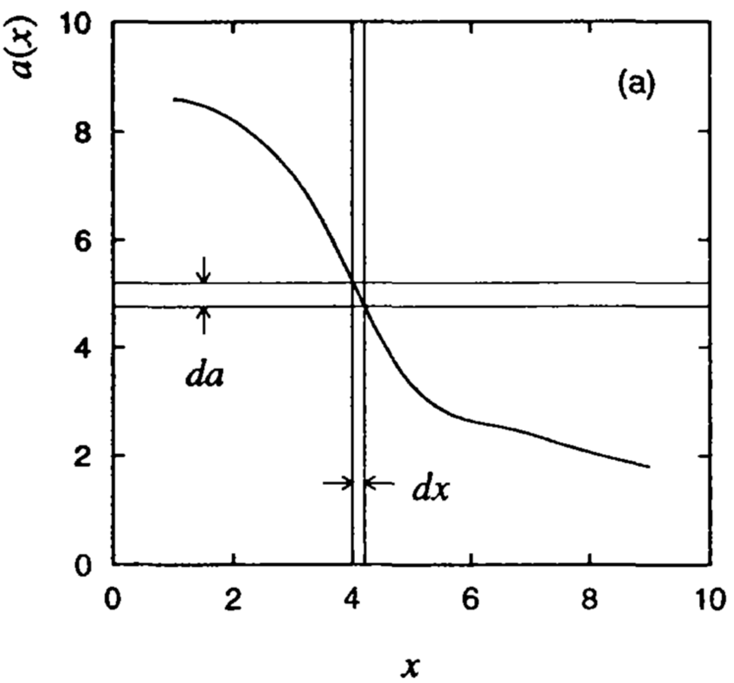
เราจะได้

$$g(a)da = f(x)dx$$

หรือ

$$\boxed{g(a) = f\big(x(a)\big)\left|\frac{dx}{da}\right|}$$

โดยเราเขียน $g(a)$ เป็น p.d.f. ของตัวแปร a และค่าสัมบูรณ์มีไว้เพื่อบังคับให้ความน่าจะเป็นเป็นบวกเสมอ



ในกรณีของหลายตัวแปรเราสามารถขยายผลนี้ไปเป็นสูตรที่ใช้จาโคเบียนในการคำนวณได้ ถ้าพิจารณาตัวแปรสุ่มจำนวน n ตัว $\vec{x} = (x_1, \dots, x_n)$ และฟังก์ชัน $a_i(\vec{x})$ โดย $i = 1, \dots, n$ เราจะได้ความสัมพันธ์ของ p.d.f. เป็น

$$g(a_1, \dots, a_n) = f(x_1, \dots, x_n) \times J$$

โดย J เป็นจาโคเบียน

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial a_1} & \frac{\partial x_1}{\partial a_2} & \cdots & \frac{\partial x_1}{\partial a_n} \\ \frac{\partial x_2}{\partial a_1} & \frac{\partial x_2}{\partial a_2} & \cdots & \frac{\partial x_2}{\partial a_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial a_1} & \frac{\partial x_n}{\partial a_2} & \cdots & \frac{\partial x_n}{\partial a_n} \end{vmatrix}$$

วิธีนี้ในการคำนวณมักจะต้องใช้วิธีการเชิงตัวเลขและมีความยุ่งยาก อย่างไรก็ดีเรามักจะสนใจแค่ค่าเฉลี่ยและความแปรปรวนของตัวแปรสุ่มซึ่งความสัมพันธ์ของตัวแปรเหล่านี้สามารถทำได้ผ่านทาง**การส่งผ่านค่าคลาดเคลื่อน (error propagation)** ซึ่งเราจะได้กล่าวถึงในส่วนถัดไป

4.5) ค่าคาดหวัง (Expectation values)

อีกค่าหนึ่งที่สำคัญในการอธิบายสถิติของข้อมูลคือ**ค่าคาดหวัง (expectation value) หรือค่าเฉลี่ยประชากร (population mean)** $E[x]$ โดย x เป็นตัวแปรสุ่ม ซึ่งนิยามโดย

$$\boxed{E[x] = \int_{-\infty}^{\infty} x\, f(x) dx = \mu}$$

โดยที่ $f(x)$ คือ p.d.f. และค่าคาดหวังของ x มักถูกเขียนแทนด้วย μ สังเกตว่าค่าคาดหวังไม่ใช่ฟังก์ชันของ x แต่ขึ้นกับ p.d.f. ตัวอย่างเช่นถ้า p.d.f. มีความหนาแน่นมากที่บริเวณเดียว ค่าคาดหวังจะมีค่าที่ x ของบริเวณนั้น หรือถ้ามีความหนาแน่นมากที่สองตำแหน่งค่าคาดหวังจะเป็นค่า x ที่อยู่ระหว่างสองตำแหน่งนั้นๆ

สำหรับฟังก์ชันของตัวแปรสุ่ม $a(x)$ ค่าคาดหวังจะเป็น

$$\boxed{E[a] = \int_{-\infty}^{\infty} a\, g(a)\, da = \int_{-\infty}^{\infty} a(x) f(x) dx}$$

โดยที่ $g(a)$ และ $f(x)$ เป็น p.d.f. ของตัวแปร a และ x ตามลำดับ

ตัวอย่างค่าคาดหวังของฟังก์ชันตัวแปรสุ่มเช่น โมเมนต์ที่ n ของ x (n^{th} moment of x)

$$E[x^n] = \int_{-\infty}^{\infty} x^n f(x) dx$$

และโมเมนต์กลางที่ n ของ x (n^{th} central moment of x)

$$E[(x - E[x])^n] = \int_{-\infty}^{\infty} (x - \mu)^n f(x) dx$$

โดยกรณีพิเศษคือโมเมนต์กลางที่ 2 ซึ่งเราเรียกว่าความแปรปรวนของประชากร (population variance)

$$E[(x - E[x])^2] = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx \equiv \sigma^2 \equiv V[x]$$

และรากที่สองของความแปรปรวนเราเรียกว่าค่าเบี่ยงเบนมาตรฐาน (standard deviation) ซึ่งอยู่ในหน่วยเดียวกันกับตัวแปรสุ่ม x

$$\sigma = \sqrt{V[x]}$$

สังเกตว่า

$$E[(x - E[x])^2] = E[x^2 - 2xE[x] + E[x]^2] = E[x^2] - 2E[x]E[x] + E[x]^2$$

$$\boxed{\sigma^2 = E[x^2] - E[x]^2}$$

สำหรับฟังก์ชันหลายตัวแปร $a(\vec{x})$ เราจะได้ค่าคาดหวังเป็น

$$E[a(\vec{x})] = \int_{-\infty}^{\infty} ag(a)da = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} a(\vec{x})f(\vec{x})dx_1 \cdots dx_n \equiv \mu_a$$

และค่าความแปรปรวนเป็น

$$V[a] = E[(a - \mu_a)] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (a(\vec{x}) - \mu_a)^2 f(\vec{x})dx_1 \cdots dx_n \equiv \sigma_a^2$$

ในกรณีของตัวแปรสุ่มสองตัวแปรเรานิยาม**ความแปรปรวนร่วม (covariance)** เป็น

$$V_{xy} = E\big[(x - \mu_x)(y - \mu_y)\big] = E\big[xy - \mu_x y - x\mu_y + \mu_x\mu_y\big] = E[xy] - \mu_x\mu_y$$

$$\boxed{V_{xy} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x,y)dxdy - \mu_x\mu_y}$$

โดย $\mu_x = E[x]$ และ $\mu_y = E[y]$ แมทริกซ์ความแปรปรวนร่วมมีอีกชื่อว่า**แมทริกซ์ความคลาดเคลื่อน (error matrix)** ซึ่งจะนิยามเป็น $cov[x,y]$ สำหรับฟังก์ชัน $a(\vec{x})$ และ $b(\vec{x})$ ของตัวแปรสุ่ม $\vec{x}$ จำนวน n ตัว แมทริกซ์ความแปรปรวนจะเขียนได้เป็น

$$cov[a,b] = E[(a - \mu_a)(b - \mu_b)] = E[ab] - \mu_a\mu_b = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} abg(a,b)dadb - \mu_a\mu_b$$

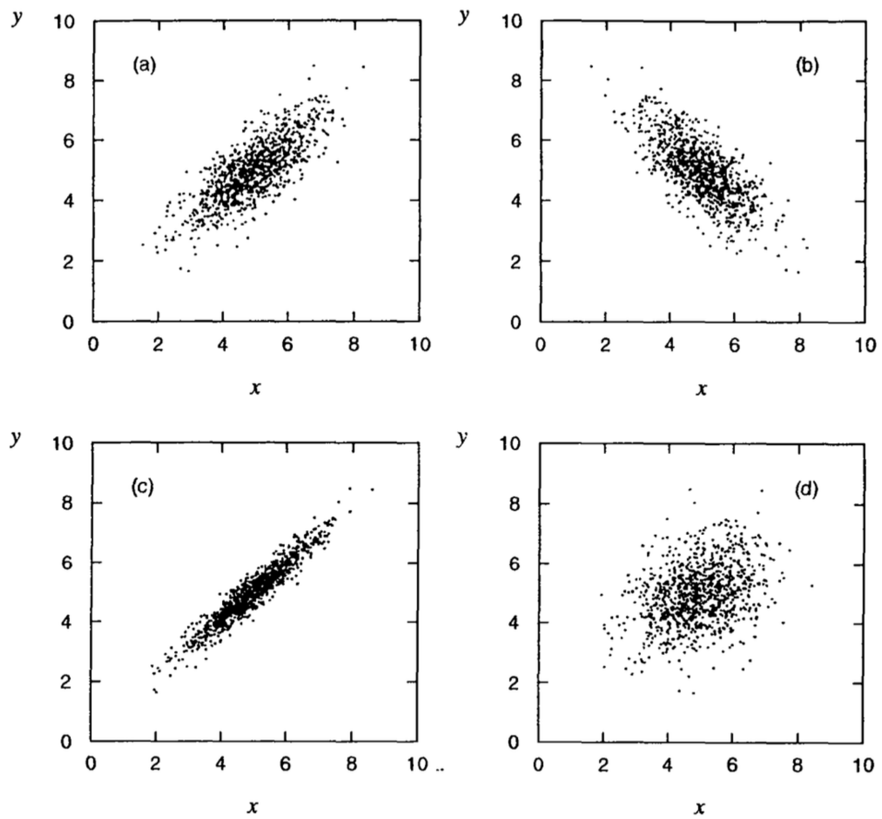
$$\boxed{cov[a,b] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} a(\vec{x})b(\vec{x})f(\vec{x})dx_1 \cdots dx_n - \mu_x\mu_y}$$

โดยที่ $g(a,b)$ เป็น p.d.f ร่วมของฟังก์ชัน $a(\vec{x})$ และ $b(\vec{x})$ สังเกตว่า $V_{ab} = cov[a,b]$ เป็นสมมาตรใน a และ b และ $V_{aa} \equiv \sigma_a^2$ (ค่าความแปรปรวนของ a) เป็นบวกเสมอ

เพื่อที่จะวัดค่าความเกี่ยวข้องกันของตัวแปรสุ่มสองตัวในรูปแบบของตัวแปรที่ไม่มีหน่วยเราจะให้ค่า**สัมประสิทธิ์สหสัมพันธ์ (correlation coefficient)**

$$\rho_{xy} = \frac{V_{xy}}{\sigma_x \sigma_y}$$

ซึ่งค่านี้มีค่าอยู่ในช่วง $-1 \leq \rho_{xy} \leq 1$



ตัวอย่างของค่าสหสัมพันธ์ (a) $\rho_{xy} = 0.75$ (b) $\rho_{xy} = -0.75$ (c) $\rho_{xy} = 0.95$ (d) $\rho_{xy} = 0.25$ กรณีของรูป (a) จะสังเกตว่าถ้า x มีค่าสูงจะมีแนวโน้มที่ y มีค่าสูงตามไปด้วย กรณีแบบนี้ซึ่งมีค่า $\rho_{xy} \sim 1$ เราจะเรียกว่ามีสหสัมพันธ์เป็นบวก (positively correlated) ในกรณีของรูป (b) สังเกตได้ว่าถ้า x มีค่าสูงจะมีแนวโน้มที่ y มีค่าต่ำลงไปอย่างสอดคล้องกัน กรณีแบบนี้ซึ่งมีค่า $\rho_{xy} \sim -1$ เราจะเรียกว่ามีสหสัมพันธ์เป็นลบ negatively correlated

ส่วนในกรณีที่ค่า $\rho_{xy} \sim 0$ จะเกิดขึ้นตอนที่ตัวแปรไม่มีสหสัมพันธ์หรือความเกี่ยวข้องกัน

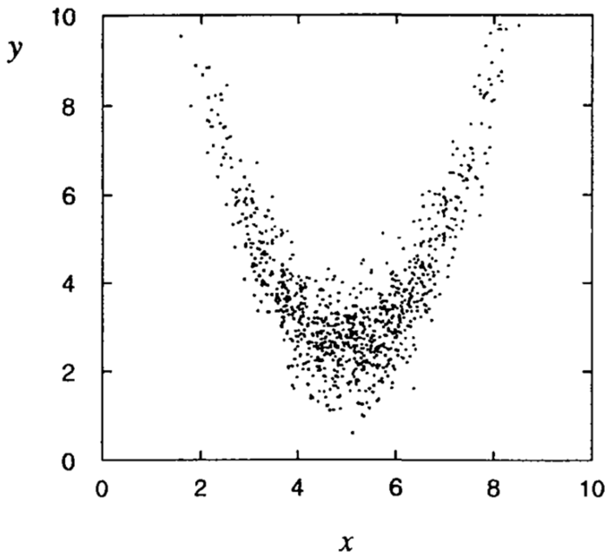
ในกรณีหนึ่งที่ทำให้ค่าสหสัมพันธ์มีค่าเป็นศูนย์คือตัวแปรสุ่มเป็นอิสระกันนั่นคือ

$$E[xy] = E[x]E[y] = \mu_x \mu_y$$

จึงจะทำให้

$$V_{xy} = E[xy] - \mu_x \mu_y = 0$$

อย่างไรก็ตามถึงแม้ว่าตัวแปรสุ่มไม่เป็นอิสระต่อกันก็ยังมีโอกาสที่ทำให้ค่าสหสัมพันธ์มีค่าเป็นศูนย์ได้เช่นภาพด้านล่างซึ่งตัวแปรสุ่มไม่เป็นอิสระต่อกัน แต่มีค่าสหสัมพันธ์เป็นบวกและเป็นลบเท่าๆ กันจึงทำให้ผลรวมเป็นศูนย์



4.5) การส่งผ่านความคลาดเคลื่อน

พิจารณาตัวแปรสุ่ม n ตัว $\vec{x} = (x_1, \dots, x_n)$ ซึ่งมี p.d.f. ร่วมเป็น $f(\vec{x})$ ซึ่งเราจะสมมติว่าเราไม่ทราบหน้าตาของ p.d.f. ร่วมนี้ แต่เราทราบถึงค่าเฉลี่ย (ค่าคาดหวัง) ของแต่ละตัวแปรสุ่ม $\vec{\mu} = (\mu_1, \dots, \mu_n)$ และเราทราบค่าแมทริกซ์ความแปรปรวน V_{ij} (จากการประมาณ) เราจะหาความแปรปรวนของฟังก์ชัน $y(\vec{x})$ ได้จากการประมาณต่อไปนี้

พิจารณาการกระจายเทย์เลอร์ของ $y(\vec{x})$ รอบๆ ค่าเฉลี่ย $\vec{x} = \vec{\mu}$

$$y(\vec{x}) \approx y(\vec{\mu}) + \sum_{i=1}^{n} \left[\frac{\partial y}{\partial x_i}\right]_{\vec{x}=\vec{\mu}} (x_i - \mu_i)$$

ดังนั้นค่าคาดหวังของ y จะประมาณได้เป็น

$$E[y(\vec{x})] \approx E[y(\vec{\mu})] + \sum_{i=1}^{n} \left[\frac{\partial y}{\partial x_i}\right]_{\vec{x}=\vec{\mu}} E[(x_i - \mu_i)] = y(\vec{\mu})$$

ส่วนค่าคาดหวังของ y^2 จะประมาณได้เป็น

$$E[y^2(\vec{x})] \approx E\left[\left(y(\vec{\mu}) + \sum_{i=1}^{n} \left[\frac{\partial y}{\partial x_i}\right]_{\vec{x}=\vec{\mu}} (x_i - \mu_i)\right)^2\right]$$

$$E[y^2(\vec{x})] \approx E\left[y^2(\vec{\mu}) + 2y(\vec{\mu}) \sum_{i=1}^{n} \left[\frac{\partial y}{\partial x_i}\right]_{\vec{x}=\vec{\mu}} (x_i - \mu_i) \right.$$

$$\left. + \left(\sum_{i=1}^{n} \left[\frac{\partial y}{\partial x_i}\right]_{\vec{x}=\vec{\mu}} (x_i - \mu_i)\right)\left(\sum_{j=1}^{n} \left[\frac{\partial y}{\partial x_j}\right]_{\vec{x}=\vec{\mu}} (x_j - \mu_j)\right)\right]$$

$$E[y^2(\vec{x})] \approx y^2(\vec{\mu}) + \sum_{i,j=1}^{n} \left[\frac{\partial y}{\partial x_i}\frac{\partial y}{\partial x_j}\right]_{\vec{x}=\vec{\mu}} V_{ij}$$

ดังนั้นค่าความแปรปรวนจะได้เป็น

$$\boxed{\sigma_a^2 = E[y^2(\vec{x})] - E[y(\vec{x})]^2 = \sum_{i,j=1}^{n} \left[\frac{\partial y}{\partial x_i}\frac{\partial y}{\partial x_j}\right]_{\vec{x}=\vec{\mu}} V_{ij}}$$

ตัวอย่าง จงพิจารณาค่าความแปรปรวนของ $y = x_1 + x_2$ ในรูปของ $\sigma_1, \sigma_2, V_{12}$

ตัวอย่าง จงพิจารณาค่าความแปรปรวนของ $y = x_1 \, x_2$ ในรูปของ $\sigma_1, \sigma_2, V_{12}$

ตัวอย่าง แท่งเหล็กแท่งหนึ่งมีความกว้าง ความยาว ความสูงเป็น $10 \; cm \times 5 \; cm \times 2 \; cm$ โดยมีค่าคลาดเคลื่อนจากการวัดของการวัดความกว้าง ความยาว ความสูงเป็น $0.05 \; cm$ จงหาความคลาดเคลื่อนของกล่องนี้เมื่อพิจารณาว่าความกว้างความยาวความสูงเป็นตัวแปรสุ่มที่อิสระต่อกัน

ตัวอย่าง พิจารณา p.d.f. ของความน่าจะเป็นที่จะพบอนุภาคในบ่อศักย์สูงอนันต์

$$\rho_n(x) = |\psi_n(x)|^2 = \begin{cases} \dfrac{2}{a}\sin^2\left(\dfrac{n\pi x}{a}\right) & 0 \le x \le a \\ 0 & \text{ที่อื่น} \end{cases}$$

1. จงหาความน่าจะเป็นที่จะพบอนุภาคในช่วง $0 \le x \le \frac{a}{2}$ ในกรณีที่ $n = 1$ และ $n = 2$
2. จงหา c.d.f. ในกรณีที่ $n = 1$ และ $n = 2$
3. จงหาค่าคาดหวังของตำแหน่ง $E[x]$ และค่าคาดหวังของตำแหน่งกำลังสอง $E[x^2]$ ในกรณีที่ $n = 1$ และ $n = 2$
4. จงหาความแปรปรวนในกรณีที่ $n = 1$ และ $n = 2$

Chapter 4: Basic Statistics

4.6) ตัวอย่างของการแจกแจงความน่าจะเป็น

ในหัวข้อนี้เราจะยกตัวอย่างการแจกแจงของ p.d.f. ที่มักจะพบได้บ่อย ซึ่งเราจะสนใจทั้งที่มาและสมบัติต่างๆเช่นพารามิเตอร์ที่ใช้อธิบายการแจกแจงเหล่านี้ (ค่าคาดหวังและความแปรปรวนหรือค่าเบี่ยงเบนมาตรฐานเป็นต้น)

การแจกแจงแบบทวินาม (binomial distribution)

พิจารณาการวัดหรือการกระทำอย่างหนึ่งซึ่งมีผลลัพธ์เพียงสองอย่างคือ “สำเร็จ” กับ “ไม่สำเร็จ” โดยที่ความน่าจะเป็นที่ผลลัพธ์จะ “สำเร็จ” คือ p เราจะมองเซตของการกระทำเหล่านี้เป็นตัวแปรสุ่มเพียงตัวแปรเดียวซึ่งแทนด้วยค่า n ที่แสดงถึงจำนวนครั้งที่สำเร็จ ดังนั้นปริภูมิตัวอย่างคือค่าทั้งหมดที่เป็นไปได้ของ n (จำนวนครั้งที่สำเร็จ) ภายในการกระทำทั้งหมด N ครั้ง เนื่องจากว่าการกระทำแต่ละครั้งเป็นอิสระซึ่งกันและกันดังนั้นความน่าจะเป็นที่การกระทำ 5 ครั้งมีผลลัพธ์เป็น “สำเร็จ สำเร็จ ไม่สำเร็จ สำเร็จ ไม่สำเร็จ” จะมีค่าเป็น

$$p \cdot p \cdot (1-p) \cdot p \cdot (1-p) = p^3(1-p)^2$$

ดังนั้นในกรณีที่ผลลัพธ์ออกมาเป็นสำเร็จ n ครั้งเราจะได้

$$p^n(1-p)^{N-n}$$

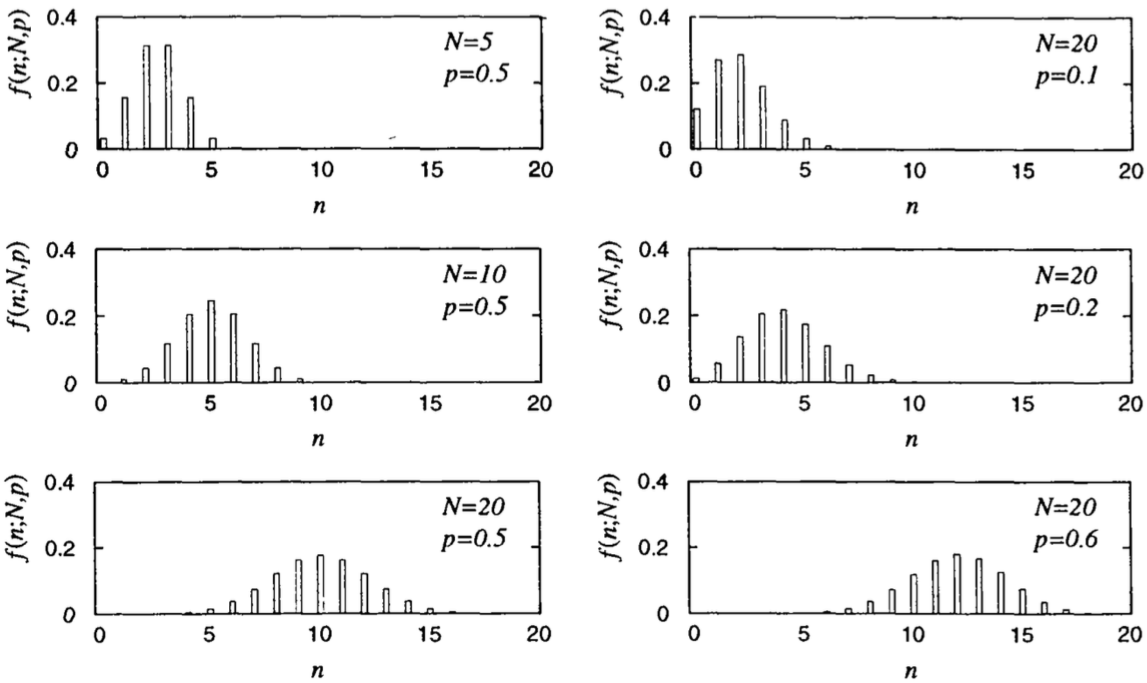
ซึ่งถ้าเราไม่สนใจในอันดับของการเกิดการสำเร็จเราจะได้ว่าจำนวนวิธีที่มีการสำเร็จเป็น n ครั้งคือ

$$\frac{N!}{n!\,(N-n)!}$$

ดังนั้นเราจะได้ความน่าจะเป็นที่ได้ผลสำเร็จ n ครั้งคือ**การแจกแจงทวินาม (binomial distribution)**

$$f(n;N,p) = \frac{N!}{n!\,(N-n)!}p^n(1-p)^{N-n}$$

โดยฟังก์ชันนี้ไม่ใช่ความหนาแน่นความน่าจะเป็นแต่บอกถึงความน่าจะเป็นซึ่งแสดงดังรูป



ค่าคาดหวังของการแจกแจงนี้มีค่าเท่ากับ

$$E[n] = \sum_{n=0}^{N} nf(n;N,p) = \sum_{n=0}^{N} n\frac{N!}{n!\,(N-n)!}p^n(1-p)^{N-n}$$

$$= \sum_{n=1}^{N} \frac{N!}{(n-1)!\,(N-n)!}p^n(1-p)^{N-n}$$

$$= \sum_{n=1}^{N} \frac{N\,(N-1)!}{(n-1)!\,\big((N-1)-(n-1)\big)!}p^n(1-p)^{N-n}$$

$$E[n] = N\sum_{n\prime=0}^{N-1}\frac{(N-1)!}{n'!\left((N-1)-n'\right)!}p^{n'+1}(1-p)^{(N-1)-n\prime}$$

$$= Np\sum_{n\prime=0}^{N\prime}\frac{N'!}{n'!\,(N'-n')!}p^{n'}(1-p)^{N\prime-n\prime} = Np$$

$$\boxed{E[n] = Np}$$

ในทำนองเดียวกันถ้าเราแสดงได้ว่า

$$E[n^2] = N^2p^2 + Np(1-p)$$

ค่าความแปรปรวนมีค่าเท่ากับ

$$\boxed{V[n] = Np(1-p)}$$

ตัวอย่าง ถ้าเราโยนเหรียญที่มีโอกาสออกหัวด้วยความน่าจะเป็น 0.5 ต่อการโยนหนึ่งครั้ง ถ้าทำการโยนเหรียญ 12 ครั้งแล้วนับจำนวนครั้งที่ออกหัว เมื่อทำการโยน 12 ครั้งแบบนี้ต่อไปเรื่อยๆ จงหาจำนวนครั้งที่ออกหัวโดยเฉลี่ย พร้อมทั้งบอกค่าเบี่ยงเบนมาตรฐาน

การแจกแจงปัวซอง (Poisson distribution)

พิจารณาการแจกแจงทวินามโดยที่ให้ N มีค่ามากๆในขณะที่ความน่าจะเป็นที่สำเร็จ p มีค่าน้อยๆ แต่ค่าคาดหวัง $Np = \lambda$ มีค่าคงที่ เราจะได้

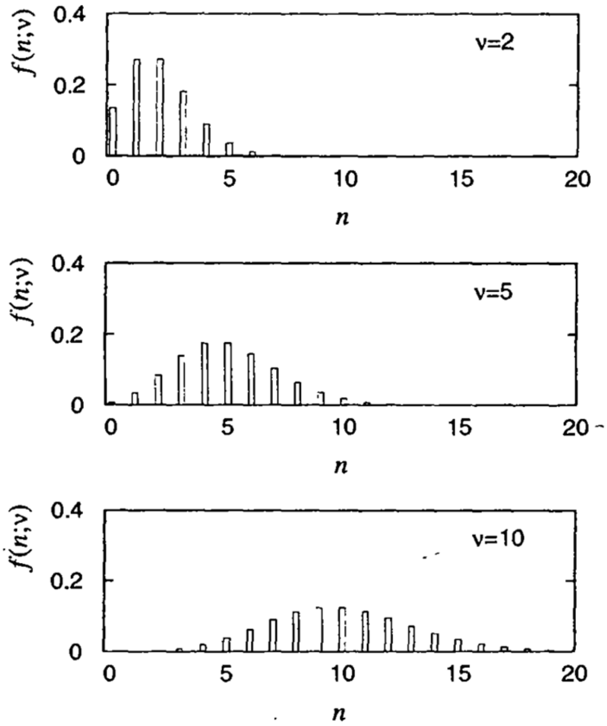
$$f(n,\lambda) = \frac{N!}{n!\,(N-n)!}\left(\frac{\lambda}{N}\right)^n\left(1-\frac{\lambda}{N}\right)^{N-n} = \frac{N(N-1)\cdots(N-n+1)}{n!}\left(\frac{\lambda}{N}\right)^n\left(1-\frac{\lambda}{N}\right)^{N-n}$$

$$= \frac{N}{N}\frac{(N-1)}{N}\cdots\frac{N-n+1}{N}\cdot\frac{\lambda^n}{n!}\left(1-\frac{\lambda}{N}\right)^N\left(1-\frac{\lambda}{N}\right)^{-n}$$

$$f(n,\lambda) = \lim_{N\to\infty}\frac{N}{N}\frac{(N-1)}{N}\cdots\frac{N-n+1}{N}\cdot\frac{\lambda^n}{n!}\left(1-\frac{\lambda}{N}\right)^N\left(1-\frac{\lambda}{N}\right)^{-n} = \frac{\lambda^n}{n!}e^{-\lambda}$$

$$\boxed{f(n,\lambda) = \frac{\lambda^n e^{-\lambda}}{n!}}$$

การแจกแจงนี้เรียกว่าการแจกแจงปัวซอง (Poisson distribution) ซึ่งขึ้นกับตัวแปร λ เพียงอย่างเดียว โดยที่ค่า $n = 0, \dots, \infty$ โดยการกระจายตัวเป็นดังรูป



โดยค่าเฉลี่ยคือ

$$E[n] = \lambda$$

และค่าคาดหวังมีค่าเท่ากับ

$$V[n] = \lim_{p\to 0} \lambda(1-p) = \lambda$$

ตัวอย่างของการแจกแจงแบบปัวซองคือจำนวนครั้งที่เกิดการสลายตัวของสารกัมมันตรังสีภายในช่วงเวลาหนึ่ง โดยในลิมิตที่จำนวนการสลายตัวที่เป็นไปได้ทั้งหมดมีค่ามากและความน่าจะเป็นที่อะตอมตัวหนึ่งจะสลายตัวในช่วงเวลาหนึ่งมีค่าน้อยมาก

ตัวอย่าง พิจารณา CS-137 ที่มีค่ากัมมันตรังสีเป็น 1 Bq ซึ่งมีการสลายตัว 10 ครั้งภายใน 10 วินาที จงหาความน่าจะเป็นที่ CS-137 มีการสลายตัว 15 และ 20 ครั้งใน 10 วินาทีตามลำดับ

การแจกแจงแบบสม่ำเสมอ (Uniform distribution)

การแจกแจงแบบสม่ำเสมอสำหรับตัวแปร x ถูกนิยามโดย p.d.f

$$f(x;\alpha,\beta) = \begin{cases} \dfrac{1}{\beta-\alpha}, & \alpha \le x \le \beta \\ 0 & \text{ที่อื่น} \end{cases}$$

นั่นคือ x มีโอกาสเท่ากันในที่จะถูกพบภายในช่วง α และ β โดยค่าเฉลี่ยและความแปรปรวนหาได้จาก

$$E[x] = \int_{\alpha}^{\beta} \frac{x}{\beta-\alpha} dx = \frac{1}{2}(\alpha+\beta)$$

$$V[x] = \int_{\alpha}^{\beta} \left[x - \frac{1}{2}(\alpha+\beta)\right]^2 \frac{1}{\beta-\alpha} dx = \frac{1}{12}(\beta-\alpha)^2$$

ตัวอย่าง การสลายตัวของอนุภาค $\pi^0 \to \gamma\gamma$ ที่มีพลังงาน E_π เราจะพบว่าพลังงานของโฟตอน E_γ มีการแจกแจงอย่างสม่ำเสมอในช่วง $E_{min} = \frac{1}{2} E_\pi \left(1 - \frac{v}{c}\right)$ ถึง $E_{max} = \frac{1}{2} E_\pi \left(1 + \frac{v}{c}\right)$ จงหาค่าคาดหวังของพลังงานของโฟตอน μ_E และค่าเบี่ยงเบนมาตรฐาน σ_E พร้อมทั้งหาความน่าจะเป็นที่โฟตอนอยู่ในช่วงพลังงาน $\mu_E - \sigma_E$ จนถึง $\mu_E + \sigma_E$

การแจกแจงแบบเอกซ์โพเนนเชียล (Exponential distribution)

ความหนาแน่นของความน่าจะเป็นแบบเอกซ์โพเนนเชียลของตัวแปร x โดย $0 \leq x \leq \infty$ ถูกนิยามโดย

$$f(x;\tau) = \frac{1}{\tau} e^{-x/\tau}$$

โดย p.d.f. นี้ถูกนิยามโดยตัวแปร τ ซึ่งเป็นค่าคาดหวังพอดีนั่นคือ

$$E[x] = \frac{1}{\tau} \int_0^{\infty} x\, e^{-\frac{x}{\tau}} dx = \tau$$

และความแปรปรวนคือ

$$V[x] = \frac{1}{\tau} \int_0^{\infty} (x-\tau)^2\, e^{-\frac{x}{\tau}} dx = \tau^2$$

ตัวอย่าง การสลายตัวของอนุภาคหนึ่งมีค่าคาดหวังของเวลาในการสลายตัว (decay time) เป็น $\tau = 5\, s$ ถ้าให้เวลาในการสลายตัวของอนุภาคนี้มี p.d.f. เป็นแบบเอกซ์โพเนนเชียล จงหาความน่าจะเป็นที่อนุภาคนี้สลายตัวภายในเวลา $t < 5$

การแจกแจงแบบเกาเซียนหรือการแจกแจงปรกติ (Gaussian distribution or Normal distribution)

ฟังก์ชัน p.d.f. แบบเกาเซียนหรือแบบปรกติของตัวแปร x โดย $-\infty \leq x \leq \infty$ ถูกนิยามโดย

$$\boxed{f(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}}$$

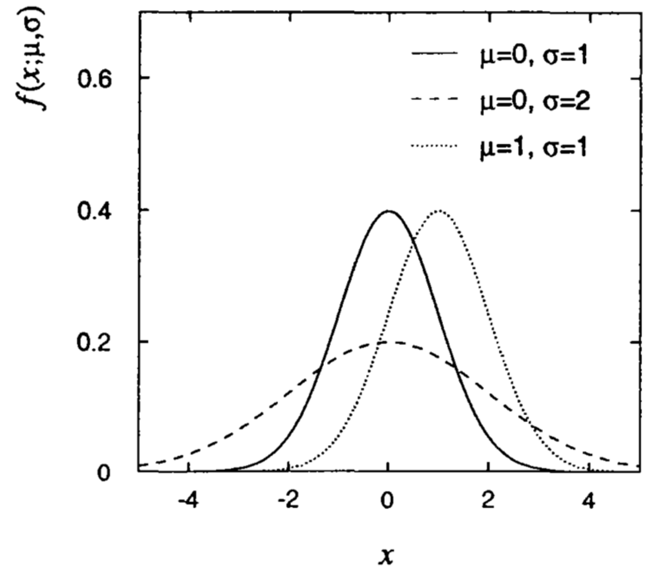
ซึ่งมีตัวแปรสองตัวเป็นตัวกำหนดฟังก์ชันนี้คือ μ และ σ ซึ่งตัวแปรนี้มีความหมายเป็นค่าคาดหวังและค่าเบี่ยงเบนมาตรฐานจากการคำนวณ

$$E[x] = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \mu$$

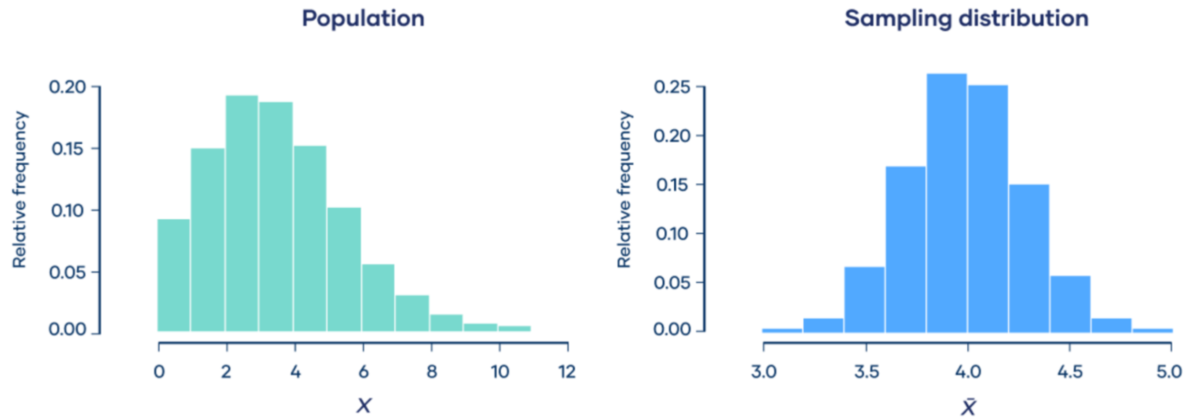
และ

$$V[x] = \int_{-\infty}^{\infty} (x-\mu)^2 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \sigma^2$$

โดยรูปร่างของ p.d.f. มีลักษณะเป็นดังรูป



ความสำคัญของการกระจายตัวแบบเกาเซียนอยู่ใน**ทฤษฎีลิมิตกลาง (Central Limit Theorem)** ซึ่งบอกเราว่าค่าเฉลี่ยของตัวแปรสุ่ม n ตัวซึ่งเป็นตัวแปรสุ่มด้วยมีการกระจายตัวแบบเกาเซียนเมื่อจำนวนตัวแปรสุ่ม n มีค่าใหญ่มากพอ ตัวอย่างเช่น ให้ตัวแปรสุ่ม X มีการกระจายตัวแบบปัวซองดังภาพซ้าย เมื่อทำการวัดตัวแปรสุ่ม X จำนวน 50 ครั้งเพื่อหาค่าเฉลี่ย $\bar{X}$ แล้วทำการวัดแบบนี้อีกหลายๆครั้ง เราจะพบว่าค่าเฉลี่ย $\bar{X}$ มีการกระจายตัวแบบเกาเซียนทางด้านขวา



ตัวอย่าง ให้ตัวแปรสุ่ม x มีการกระจายตัวแบบเกาเซียน จงหาค่าคาดหวังของ $E[x^4]$

ตัวอย่าง ถ้าให้คะแนนของนักเรียนห้องหนึ่งมีการกระจายตัวแบบเกาเซียนที่มีค่าคาดหวังที่ 50 คะแนน และค่าเบี่ยงเบนมาตรฐานเป็น 10 ถ้าข้อตกลงของวิชาบอกว่าถ้าได้คะแนนในช่วง 10% แรกของห้องจะได้เกรด A จงหาว่าคะแนนของนักศึกษาต้องมากกว่าเท่าใดจึงจะได้เกรด A

แบบฝึกหัด 4

1). สอบใบขับขี่

การสอบใบขับขี่ประกอบไปด้วยการสอบภาคทฤษฎีและการสอบภาคปฏิบัติ ถ้า 25% ของผู้เข้าสอบสอบตกภาคปฏิบัติ และ 15% ของผู้เข้าสอบสอบตกภาคทฤษฎี และ 10% ของผู้เข้าสอบสอบตกทั้งภาคทฤษฎีและภาคปฏิบัติ ถ้าเลือกสุ่มผู้เข้าสอบมาหนึ่งคน จงหาความน่าจะเป็นที่คนนี้จะ

1.1) สอบตกอย่างน้อยหนึ่งตอน

1.2) สอบตกภาคปฏิบัติ แต่สอบผ่านภาคทฤษฎี

1.3) สอบผ่านทั้งสองตอน

1.4) สอบผ่านอย่างน้อยหนึ่งตอน

2). เชื้อโรคในวัว

ถ้ามีเชื้อโรคในวัวชนิดหนึ่งที่สามารถติดต่อกันระหว่างสัตว์หลายๆประเภท สมมติให้วัว 200 ตัว ถูกทดสอบว่าติดเชื้อโรคหรือไม่ โดยเหตุการณ์ A คือเหตุการณ์ที่วัวถูกขนส่งด้วยรถบรรทุกเมื่อไม่นานมานี้ และเหตุการณ์ B คือเหตุการณ์ที่วัวถูกทดสอบเชื้อโรคแล้วมีผลเป็นบวก (ติดโรค) โดยข้อมูลของการตรวจเป็นไปตามตารางด้านล่าง

	B	B'
A	40	60
A'	20	80

2.1) มีความน่าจะเป็นเท่าใดที่วัวจะติดเชื้อและถูกขนส่งด้วยรถบรรทุกเมื่อไม่นานมานี้

2.2) มีความน่าจะเป็นเท่าใดที่วัวจะติดเชื้อถ้ากำหนดให้วัวที่ทำการทดสอบนั้นถูกขนส่งด้วยรถบรรทุกเมื่อไม่นานมานี้

2.3) ความน่าจะเป็นที่วัวจะติดเชื้อ

3). คุณภาพสินค้า

กำหนดให้ x เป็นดัชนีคุณภาพที่ผู้เชี่ยวชาญให้กับสินค้า พิจารณาคุณภาพของมือถือรุ่นหนึ่งมี p.d.f. ของดัชนีคุณภาพเป็น

$$f(x) = \begin{cases} cx(2-x), & 0 \leq x \leq 2 \\ 0, & \text{ที่อื่น} \end{cases}$$

3.1) จงหาค่าของ c

3.2) จงหา c.d.f.

3.3) จงคำนวณหาค่าคาดหวังและความแปรปรวน $E[x]$, $V[x]$

3.4) มีความน่าจะเป็นเท่าใดที่มือถือเครื่องหนึ่งมีดัชนีคุณภาพมากกว่า 1 ($x > 1$)

4). Probability density function

จาก cumulative distribution function (c.d.f)

$$F(x) = \begin{cases} 0, & x < 0 \\ 3x^2 - 2x^3, & 0 \leq x \leq 1 \\ 1, & x > 1 \end{cases}$$

4.1) จงหาฟังก์ชัน p.d.f.

4.2) จงหาความน่าจะเป็นที่ $x < 3$

4.3) จงคำนวณหาค่าคาดหวังและความแปรปรวน $E[x]$, $V[x]$

5). Binomial distribution

บริษัทผลิตขนมถุงทำการฉลองครบรอบ 50 ปีด้วยการใส่ของเล่นในถุงขนมทุกๆ 6 ถุง (ความน่าจะเป็นที่พบของเล่นในถุงคือ 1/6) ถ้าเราต้องการของเล่นนี้และทำการซื้อขนมมาทั้งหมด 20 ถุง

5.1) จงหาความน่าจะเป็นที่จะพบของเล่น 4 ชิ้นจากจำนวน 20 ถุง

5.2) จงหาความน่าจะเป็นที่ไม่พบของเล่นเลยจากจำนวน 20 จุง

6). Poisson distribution

บริษัทประกันแห่งหนึ่งกำลังพิจารณาการออกประกันสำหรับภัยธรรมชาติ จากประสบการณ์แล้วเราพบว่าจำนวนภัยธรรมชาติในช่วงหน้าหนาว W เป็นตัวแปรสุ่มที่มีการกระจายตัวแบบปัวซองที่มี $\lambda_W = 4$ และ จำนวนภัยธรรมชาติในช่วงหน้าร้อน S เป็นตัวแปรสุ่มที่มีการกระจายตัวแบบปัวซองที่มี $\lambda_S = 3$ จงหาความน่าจะเป็นที่เกิดภัยธรรมชาติในช่วงหน้าหนาวอย่างน้อยหนึ่งครั้งและเกิดภัยธรรมชาติในช่วงหน้าร้อนอย่างน้อยหนึ่งครั้ง (สมมติว่าตัวแปรสุ่มทั้งสองเป็นอิสระต่อกัน)

7). Uniform distribution

กำหนดให้เวลาที่นักเรียนคนหนึ่งทำข้อสอบวิชา Mathematical Method in Physics II มีการกระจายตัวแบบสม่ำเสมอ (uniform) ระหว่าง 90 นาที ถึง 240 นาที

7.1) จงหาความน่าจะเป็นที่นักเรียนคนหนึ่งจะใช้เวลามากกว่า 120 นาทีในการทำข้อสอบ

7.2) ถ้านักเรียนคนหนึ่งไม่ออกจากห้องสอบหลังจาก 120 นาที จงหาความน่าจะเป็นที่นักเรียนคนนี้ใช้เวลาทำข้อสอบมากกว่า 200 นาที

8). Exponential distribution

กำหนดให้จำนวนวันที่นักท่องเที่ยวจองตั๋วเครื่องบินก่อนถึงวันเดินทางมีการกระจายตัวแบบ exponential distribution โดยมีค่าเฉลี่ยอยู่ที่ 15 วัน

8.1) จงหาความน่าจะเป็นที่นักท่องเที่ยวคนหนึ่งจะซื้อตั๋วล่วงหน้าน้อยกว่า 10 วัน

8.2) จงหาว่า quantile ที่ p = 0.5 มีค่าเท่ากับกี่วัน

9). Gaussian distribution

จากการศึกษาขนาดของไข่นกชนิดหนึ่งเราพบว่า ขนาดไข่มีความยาวเป็นตัวแปรสุ่มที่มีการกระจายตัวแบบเกาเซียนด้วย $\mu = 42.1$ mm และ $\sigma = 20.8$ mm

9.1) จงหาความน่าจะเป็นที่ไข่มีความยาวมากกว่า 50 mm

9.2) จงหาความน่าจะเป็นที่ไข่มีความยาวระหว่าง 30 mm และ 40 mm

Chapter 5: Parameter Estimation and Hypothesis Testing

หลังจากที่เราได้รู้จักความน่าจะเป็นและค่าทางสถิติในเบื้องต้นแล้ว เราจะทำการเชื่อมโยงกับผลการสังเกตการณ์ (การวัด) โดยปัญหาหลักของทางสถิติคือความพยายามในการหา p.d.f. ของตัวแปรสุ่มจากการสังเกตการณ์

5.1) การประมาณค่าพารามิเตอร์ (Parameter Estimation)

พิจารณาตัวแปรสุ่ม x ซึ่งมี p.d.f. เป็น $f(x)$ ในกรณีนี้ปริภูมิตัวอย่าง (sample space) คือค่าที่เป็นไปได้ทั้งหมดของ x ถ้าเราทำการสังเกตตัวแปรสุ่มนี้แบบซ้ำๆจำนวน n ครั้ง เซตของการสังเกตการณ์ของตัวแปรสุ่ม x จำนวน n ครั้งนี้จะเรียกว่า**ตัวอย่างขนาด n (sample of size n)** เนื่องจากตัวอย่างขนาด n นี้สามารถคิดเป็นตัวแปรสุ่ม $\vec{x} = (x_1, \dots, x_n)$ ได้ ดังนั้นปริภูมิตัวอย่างของตัวอย่างขนาด n จึงเป็นค่าของเวกเตอร์ $\vec{x}$ ที่เป็นไปได้ทั้งหมด เนื่องจากการวัดแต่ละครั้งนั้นเป็นอิสระต่อกันและมี p.d.f. เหมือนกันหมดคือ $f(x)$ ดังนั้น p.d.f. ร่วมของตัวอย่างขนาด n คือ

$$f_{sample}(x_1, \dots, x_n) = f(x_1)f(x_2)\cdots f(x_n)$$

ในกรณีนี้เราจะเรียกตัวแปรสุ่ม $\vec{x} = (x_1, \dots, x_n)$ ว่าเป็นตัวแปรสุ่มที่**อิสระและกระจายตัวเหมือนกัน (independent and identically distributed, i.i.d.)** ถึงแม้ว่าจำนวนครั้งในการวัด (ขนาดตัวอย่าง) จะมีค่ามากแต่รูปแบบของ p.d.f. ร่วมนี้ก็มีความเรียบง่ายเพราะคือการคูณกันของ p.d.f. ที่เหมือนกันทุกประการ

พิจารณาถึงสถานการณ์ที่ p.d.f. ของการวัดหนึ่งครั้งขึ้นกับพารามิเตอร์ $\vec{\theta} = (\theta_1, \dots, \theta_m)$ นั่นคือ p.d.f. เป็น $f(x;\vec{\theta})$ ตัวอย่างเช่นค่า λ ใน Poisson distribution, ค่า τ ใน Exponential distribution หรือค่า μ, σ ใน Gaussian distribution เป็นต้น เป้าหมายของเราในบทนี้คือการประมาณค่าพารามิเตอร์ $\vec{\theta} = (\theta_1, \dots, \theta_m)$ จากการวัด $\vec{x} = (x_1, \dots, x_n)$ ที่เรามี

โดยฟังก์ชันที่เราสร้างจากการวัด $\vec{x} = (x_1, \dots, x_n)$ จะเรียกว่า**สถิติ (statistic)** และค่าสถิติที่ใช้ในการประมาณค่าเราจะเรียกว่า**ตัวประมาณค่า (estimator)** โดยปรกติตัวประมาณค่าเราจะเขียนด้วยสัญลักษณ์ $\hat{\theta}$ เพื่อที่จะแยกความแตกต่างจากพารามิเตอร์จริง θ

ถ้า $\hat{\theta}$ ลู่เข้าหา θ ในลิมิตที่ n มีค่ามาก ตัวประมาณค่านี้จะเรียกว่า**ตัวประมาณที่สอดคล้องกัน (consistent)** ซึ่งนี่เป็น<u>เงื่อนไขที่จำเป็นของการเป็นตัวประมาณค่า</u> (ถ้าไม่ลู่เข้าหาค่าจริงก็จะเป็นตัวประมาณค่าที่ไม่ดี) กระบวนการประมาณค่าพารามิเตอร์จากการวัดจะถูกเรียกว่าการปรับพารามิเตอร์ให้เหมาะสม (parameter fitting) เนื่องจาก $\hat{\theta}$ เป็นฟังก์ชันของตัวแปรสุ่ม ดังนั้นมันก็เป็นตัวแปรสุ่มเช่นเดียวกัน โดยเราคาดหวังว่า $E[\hat{\theta}] \to \theta$ อย่างไรก็ดีในบางครั้งตัวประมาณค่าที่สอดคล้องกันก็ไม่ได้มีสมบัติข้างต้น ดังนั้นเราจะนิยาม**ไบแอส (bias)** เป็น

$$b = E[\hat{\theta}] - \theta$$

โดยตัวประมาณที่สอดคล้องกันไม่จำเป็นต้องมีไบแอสเป็นศูนย์ อีกปริมาณหนึ่งที่สำคัญคือค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Squared Error, MSE) ซึ่งนิยามเป็น

$$MSE = E\left[\left(\hat{\theta} - \theta\right)^2\right] = E\left[\left(\hat{\theta} - E[\hat{\theta}] + E[\hat{\theta}] - \theta\right)^2\right]$$

$$= E\left[\left(\hat{\theta} - E[\hat{\theta}]\right)^2\right] + 2E\left[\left(\hat{\theta} - E[\hat{\theta}]\right)\left(E[\hat{\theta}] - \theta\right)\right] + E\left[\left(E[\hat{\theta}] - \theta\right)^2\right]$$

$$= E\left[\left(\hat{\theta} - E[\hat{\theta}]\right)^2\right] + \left(E[\hat{\theta}] - \theta\right)^2$$

$$\boxed{MSE = V[\hat{\theta}] + b^2}$$

ซึ่งเรามักจะตีความเป็นผลรวมกำลังสองของค่าคลาดเคลื่อนทางสถิติ (statistical error) และค่าคลาดเคลื่อนเชิงระบบ (systematic error)

5.2) ตัวประมาณสำหรับค่าเฉลี่ย ความแปรปรวน

ตัวอย่างที่สำคัญของตัวประมาณคือ**ค่าเฉลี่ยตัวอย่าง (sample mean)**

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

ซึ่งเราสามารถตรวจสอบได้ว่าตัวประมาณนี้ไม่มีไบแอส

$$E[\bar{x}] = \frac{1}{n}E\left[\sum_{i=1}^{n} x_i\right] = \frac{1}{n}\sum_{i=1}^{n} E[x_i] = \frac{1}{n}\sum_{i=1}^{n} \mu = \mu$$

นั่นคือ

$$E[\bar{x}] - \mu = 0$$

และความแปรปรวนมีค่าเป็น

$$V[\bar{x}] = E[(\bar{x}-\mu)^2] = E[\bar{x}^2] - E[\bar{x}]^2 = E[\bar{x}^2] - \mu^2$$

โดยที่

$$E[\bar{x}^2] = E\left[\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)\left(\frac{1}{n}\sum_{j=1}^{n} x_j\right)\right] = \frac{1}{n^2}E\left[\sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j\right] = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} E[x_i x_j]$$

$$= \frac{1}{n^2}\left[\sum_{i=j}^{n} E[x_i^2] + \sum_{i=1}^{n}\sum_{j\neq i}^{n} E\left[x_i x_j\right]\right] = \frac{1}{n^2}[n(\sigma^2+\mu^2) + (n^2-n)\mu^2] = \frac{\sigma^2}{n} + \mu^2$$

ในบรรทัดสุดท้ายเราใช้ $E\left[x_i x_j\right] = E[x_i]E\left[x_j\right] = \mu^2$ และ $\sigma^2 = E[x_i^2] - \mu^2$

และความแปรปรวนเป็น

$$V[\bar{x}] = E[\bar{x}^2] - \mu^2 = \frac{\sigma^2}{n} + \mu^2 - \mu^2 = \frac{\sigma^2}{n}$$

ซึ่งผลลัพท์นี้บอกเราว่าถ้าอยากได้ค่าเฉลี่ยตัวอย่างที่มีการกระจายตัวน้อยเราต้องเลือกกลุ่มตัวอย่างที่มีขนาดใหญ่เพราะ $n \to \infty, V[\bar{x}] \to 0$

อีกตัวประมาณหนึ่งที่สำคัญคือ**ความแปรปรวนตัวอย่าง (sample variance)**

$$\boxed{S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

ซึ่งเราสามารถตรวจสอบได้ว่าตัวประมาณนี้ไม่มีไบแอส

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \mu + \mu - \bar{x})^2 = \frac{1}{n-1}\sum_{i=1}^{n}[(x_i-\mu)^2 + 2(x_i-\mu)(\mu-\bar{x}) + (\mu-\bar{x})^2\,]$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\mu)^2 + \frac{2}{n-1}\sum_{i=1}^{n}(x_i-\mu)(\mu-\bar{x}) + \frac{1}{n-1}\sum_{i=1}^{n}(\bar{x}-\mu)^2$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\mu)^2 + \frac{2}{n-1}(n\bar{x}-n\mu)(\mu-\bar{x}) + \frac{1}{n-1}\sum_{i=1}^{n}(\bar{x}-\mu)^2$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\mu)^2 - \frac{2n}{n-1}(\bar{x}-\mu)^2 + \frac{n}{n-1}(\bar{x}-\mu)^2$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\mu)^2 - \frac{n}{n-1}(\bar{x}-\mu)^2$$

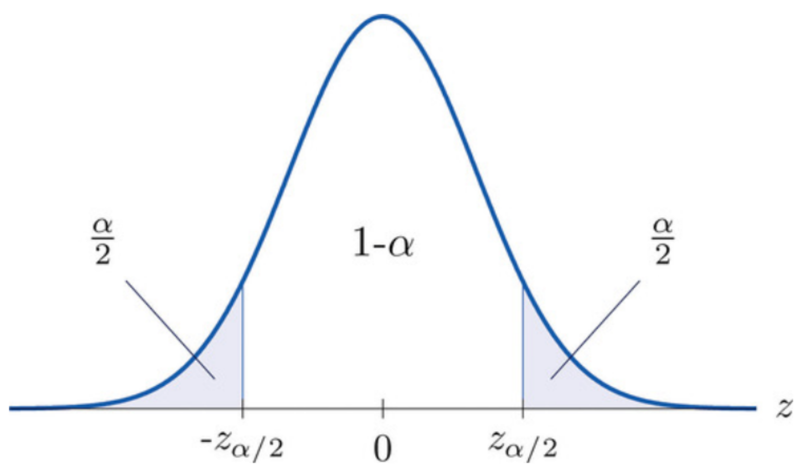
$$E[S^2] = \frac{1}{n-1}\sum_{i=1}^{n}E[(x_i-\mu)^2] - \frac{n}{n-1}E[(\bar{x}-\mu)^2] = \frac{n\sigma^2}{n-1} - \frac{n}{n-1}\frac{\sigma^2}{n} = \sigma^2$$

5.3) ระดับความเชื่อมั่น (Confidence level)

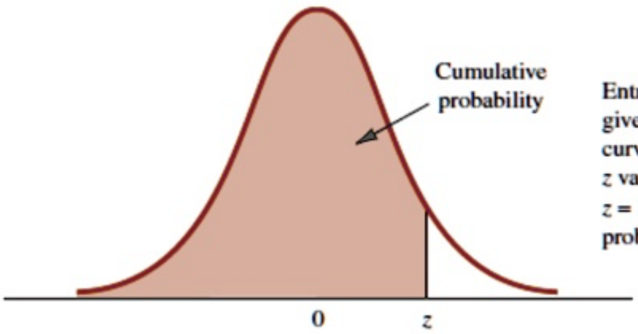
จากการใช้ค่าเฉลี่ยตัวอย่างในตอนที่แล้ว เราบอกได้ว่าค่าเฉลี่ยตัวอย่างมีค่าคาดหวังอยู่ที่ค่าคาดหวังของตัวแปรสุ่ม $E[\bar{x}] = \mu$ และความแปรปรวนมีค่าเป็น $V[\bar{x}] = \frac{\sigma^2}{n}$ นั่นคือมีค่าเบี่ยงเบนมาตรฐานเป็น $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ ซึ่งโดยปรกติเรามักจะต้องรายงานช่วงของค่า x ที่เรามั่นใจได้ในเปอร์เซ็นต์ระดับหนึ่งว่าค่า μ อยู่ภายในช่วงนี้ ช่วงที่ครอบคลุมนี้เราเรียกว่าช่วงความเชื่อมั่น (confidence interval) ตัวอย่างเช่นถ้าต้องการให้ 95% ของการสุ่มกลุ่มตัวอย่างจะมีค่าคาดหวังจริงอยู่ในช่วง $[\bar{x} - E, \bar{x} + E]$ โดยที่ค่า E คือค่าขอบเขตความคลาดเคลื่อนของการประมาณ (margin of error of the estimate) ซึ่งเรามักจะเขียนในรูปของจำนวนเท่าของ $\frac{\sigma}{\sqrt{n}}$ นั่นคือ

$$\boxed{E = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}}$$

โดยที่ค่า α คือค่าระดับความเชื่อมั่นเช่น 95% จะมีค่า α เป็น $1 - 0.95 = 0.05$ เป็นต้น
จากนิยาม ระดับความเชื่อมั่น α จะสอดคล้องกับค่า $z_{\alpha/2}$ ที่ทำให้พื้นที่ใต้กราฟมีค่าเท่ากับ $1 - \alpha$ นั่นคือสอดคล้องกับรูป



โดยปรกติเพื่อที่จะหาค่า $z_{\alpha/2}$ นี้เรามักจะใช้ตาราง



Entries in the table give the area under the curve to the left of the z value. For example, for $z = 1.25$, the cumulative probability is .8944.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990

ตัวอย่าง จงหาค่า $z_{\alpha/2}$ ของระดับความเชื่อมั่น 90%, 95%, และ 99% ตามลำดับ

ตัวอย่าง กลุ่มตัวอย่างของนักศึกษาจำนวน 120 คน มีค่าเฉลี่ยตัวอย่างของเกรดเฉลี่ยเท่ากับ 2.71 และมีค่าเบี่ยงเบนมาตรฐานตัวอย่างเท่ากับ 0.51 จงสร้างช่วงความเชื่อมั่น 99% ของเกรดเฉลี่ยนี้

ตัวอย่าง ในการทดลองลูกตุ้มเพนดูลัมอย่างง่าย เมื่อทำการวัดเวลาที่ใช้ในการเคลื่อนที่หนึ่งคาบ โดยทำการวัดซ้ำ 10 ครั้ง พบว่าค่าที่วัดได้คือ 3.6, 3.5, 3.8, 3.7, 3.4, 3.5, 3.6, 3.8, 3.1, 3.9 ตามลำดับ จงหาค่าเฉลี่ยตัวอย่าง ส่วนเบี่ยงเบนมาตรฐานตัวอย่างและช่วงความเชื่อมัน 90% ของคาบนี้

ตัวอย่าง ในการวัดความกว้างของใบไม้ โดยการวัดซ้ำทั้งหมด 10 ครั้งพบว่าความกว้างของใบไม้มีค่าเป็น 6.20, 6.40, 6.10, 6.30, 6.55, 6.05, 6.45, 6.30, 6.35, 6.60 ตามลำดับ จงหาค่าเฉลี่ยตัวอย่าง ส่วนเบี่ยงเบนมาตรฐานตัวอย่างและช่วงความเชื่อมัน 95% ของคาบนี้

Chapter 5: Parameter Estimation and Hypothesis Testing

ในตอนที่แล้วเราได้พูดถึงการประมาณค่าพารามิเตอร์ต่างๆของ p.d.f. ของตัวแปรสุ่ม และรายงานระดับความเชื่อมั่นของการประมาณนั้น อย่างไรก็ดีในการประยุกต์ใช้จริง เรามักจะสนใจเกี่ยวกับสมมติฐานมากกว่า ตัวอย่างเช่นบริษัทประกันอาจจะสนใจมากกว่าว่าอัตราการเกิดอุบัติเหตุมีค่ามากกว่า 2 ครั้งต่อปีหรือไม่ เพื่อที่จะตรวจสอบ/ทดสอบสมมติฐานนี้เราต้องใช้สถิติเพื่อการทดสอบ

5.4) การทดสอบสมมติฐาน (Hypothesis testing)

เราจะเริ่มจากการสร้างสมมติฐาน อย่างที่เคยได้บอกไว้ว่าสมมติฐานคือข้อความ/ประโยค/คำแถลงที่มีค่าเป็นจริงหรือเท็จ โดยปรกติแล้วมักจะเกี่ยวกับพารามิเตอร์ของ p.d.f. โดยเราจะแบ่งสมมติฐานออกเป็นสองประเภทคือ

1. สมมติฐานหลัก (Null hypothesis) ซึ่งคือสมมติฐานเพื่อใช้ในการทดสอบและจะถูกสมมติให้เป็นจริงจนกว่าจะมีหลักฐานที่แน่นหนามาหักล้าง มักจะเขียนแทนด้วย H_0
2. สมมติฐานแย้ง (Alternative hypothesis) ซึ่งคือสมมติฐานที่ตรงข้ามกับสมมติฐานหลักและจะให้ค่าเป็นจริงได้ก็ต่อเมื่อมีหลักฐานที่แน่นหนามาสนับสนุน มักจะเขียนแทนด้วย H_a

โดยผลของการทดสอบมีสองทาง

1. ปฏิเสธสมมติฐานหลัก H_0 (แปลว่ายอมรับสมมติฐานแย้ง H_a)
2. ไม่สามารถปฏิเสธสมมติฐานหลักได้ (และไม่สามารถยอมรับสมมติฐานแย้งได้)

ตัวอย่างของการตั้งสมมติฐาน

1. โรงอาหารมหาวิทยาลัยแห่งหนึ่งรายงานว่าค่าเฉลี่ยของอาหารต่อจานมีค่าเท่ากับ 30 บาท อย่างไรก็ตามสโมสรนักศึกษาเชื่อว่าราคาเฉลี่ยมีค่าสูงกว่านี้มาก สมมติฐานที่ตั้งจะเป็น

สมมติฐานหลัก $H_0 : \mu = 30$

สมมติฐานแย้ง $H_a : \mu > 30$

2. โรงงานผลิตขนมแห่งหนึ่งรายงานปริมาณไขมันในขนมเป็น 8 กรัมต่อชิ้น หน่วยงานควบคุมคุณภาพได้เข้าทำการตรวจสอบว่าการผลิตนี้ได้มาตรฐานตรงตามค่าที่รายงานไว้หรือไม่

สมมติฐานที่ตั้งจะเป็น

สมมติฐานหลัก $H_0 : \mu = 8$

สมมติฐานแย้ง $H_a : \mu \neq 8$

ข้อสังเกตของการตั้งสมมติฐานหลักคือเราไม่ใช้ $H_0 : \mu \geq 30$ สำหรับข้อแรก ถึงแม้ว่ามันจะดูเป็นธรรมชาติมากกว่า เหตุผลคือเพื่อที่จะทำให้การปฏิเสธสมมติฐานหลักทำได้โดยง่ายและเข้าใจง่ายสำหรับผู้ทำการทดสอบ จึงเป็นข้อตกลงว่าเราจะใช้เครื่องหมายเท่ากับสำหรับสมมุติฐานหลักเสมอ

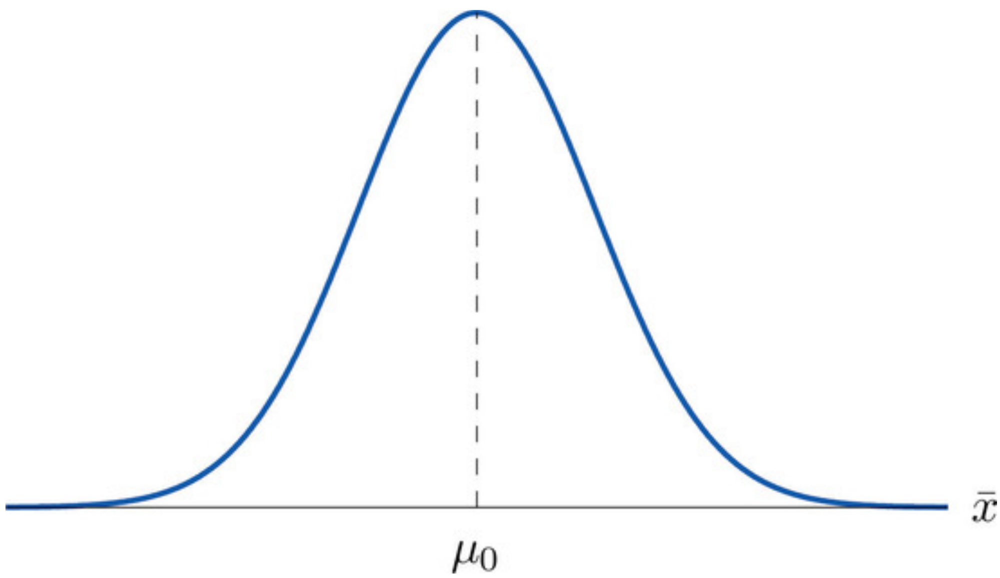
<u>วิธีการตรวจสอบสมมติฐานที่เกี่ยวข้องกับค่าเฉลี่ยประชากร (ค่าคาดหวัง)</u>

เราจะทำการหาค่าเฉลี่ยจากกลุ่มตัวอย่าง n เพื่อหาค่า $\bar{x}$ แล้วพิจารณาว่าค่า $\bar{x}$ ทีได้มีความเป็นไปได้มากแค่ไหนถ้าสมมติให้สมมติฐานหลักเป็นจริง

ถ้าสมมติฐานหลัก

$$H_0 : \mu = \mu_0$$

เป็นจริงและเราทำการสุ่มตัวอย่างที่มีขนาดใหญ่พอเราจะได้ $\bar{x}$ มี p.d.f. เป็น Gaussian distribution ซึ่งมีค่าคาดหวังอยู่ที่ μ_0 และส่วนเบี่ยงเบนมาตรฐานเป็น $\frac{\sigma}{\sqrt{n}}$ ตามภาพ

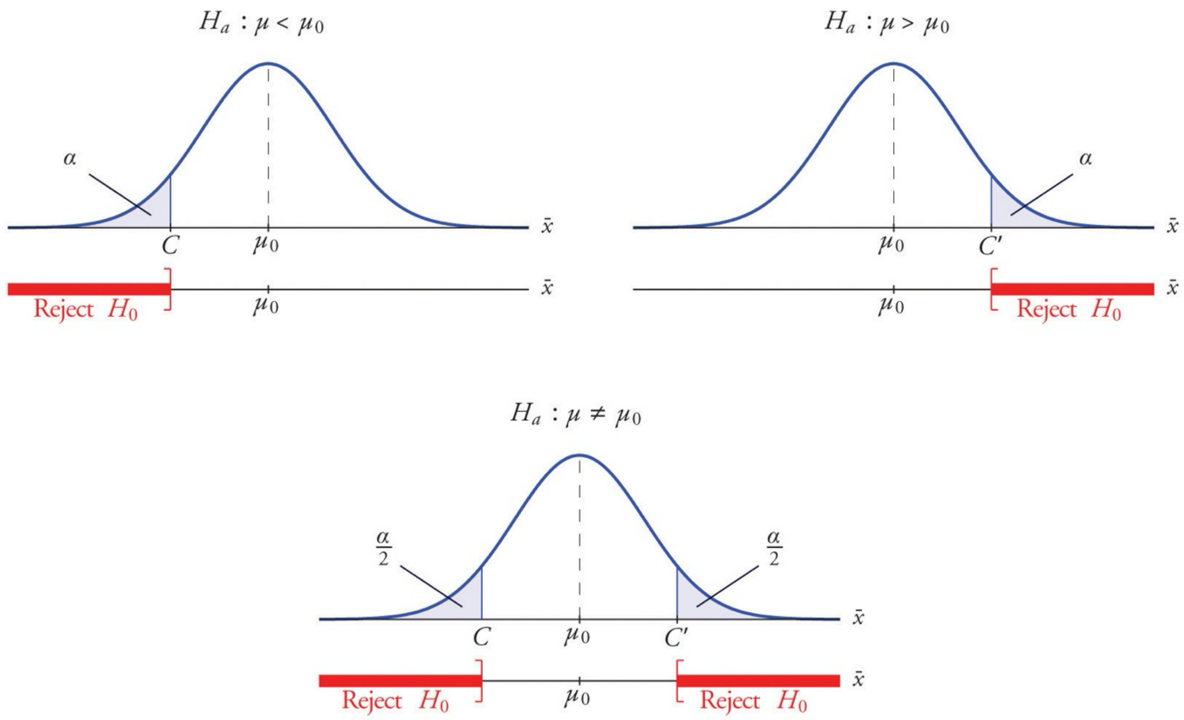


ดังนั้นวิธีการพิจารณาเพื่อปฏิเสธสมมติฐานหลักทำได้จากการดูว่าค่า $\bar{x}$ ที่มีความน่าจะเป็นเท่าใดที่จะมาจาก p.d.f. ที่สอดคล้องกับสมมติฐานหลักนี้

โดยเราจะแบ่งสมมติฐานแย้งเป็นสามประเภท

1. ถ้า $H_a : \mu < \mu_0$ เราจะปฏิเสธ H_0 เมื่อ $\bar{x}$ อยู่ห่างจาก μ_0 ไปทางซ้าย
2. ถ้า $H_a : \mu > \mu_0$ เราจะปฏิเสธ H_0 เมื่อ $\bar{x}$ อยู่ห่างจาก μ_0 ไปทางขวา
3. ถ้า $H_a : \mu \neq \mu_0$ เราจะปฏิเสธ H_0 เมื่อ $\bar{x}$ อยู่ห่างจาก μ_0 ไปทางซ้ายหรือทางขวา

โดยเราจะเลือกค่า C ค่าหนึ่งเพื่อที่จะเลือก**บริเวณปฏิเสธ (rejection region)** โดยค่านี้เราจะเลือกจากพื้นที่ใต้กราฟ α ที่แสดงถึงความน่าจะเป็นที่เหตุการณ์นี้เกิดขึ้น โดยค่า α เราจะเรียกว่า **ระดับนัยสำคัญ (significant level)**



เราจะลองพิจารณาตัวอย่างที่สอง โดยโรงงานผลิตขนมแห่งหนึ่งรายงานปริมาณไขมันในขนมเป็น 8 กรัมต่อชิ้นและมีค่าเบี่ยงเบนมาตรฐานเป็น 0.15 กรัม และมีสมมติฐานเป็น

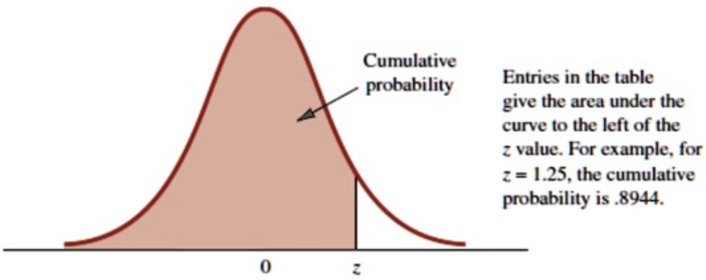
สมมติฐานหลัก $H_0 : \mu = 8$

สมมติฐานแย้ง $H_a : \mu \neq 8$

ถ้ากำหนดให้เราทำการสุ่มตัวอย่างขนาด 5 (ขนม 5 ซอง) เราสามารถสร้างบริเวณปฏิเสธที่มีค่า $\alpha = 0.10$ ได้จากการที่พิจารณาค่าคาดหวังและค่าเบี่ยงเบนมาตรฐานของ $\bar{x}$ ซึ่งจากการประมาณค่าเราพบว่า

$$E[\bar{x}] = \mu_0, \qquad \sqrt{V[\bar{x}]} = \frac{\sigma}{\sqrt{n}}$$

ดังนั้นเราต้องหาค่า z_α ที่สอดคล้องกับพื้นที่ใต้กราฟ $\alpha/2 = 0.10/2 = 0.05$



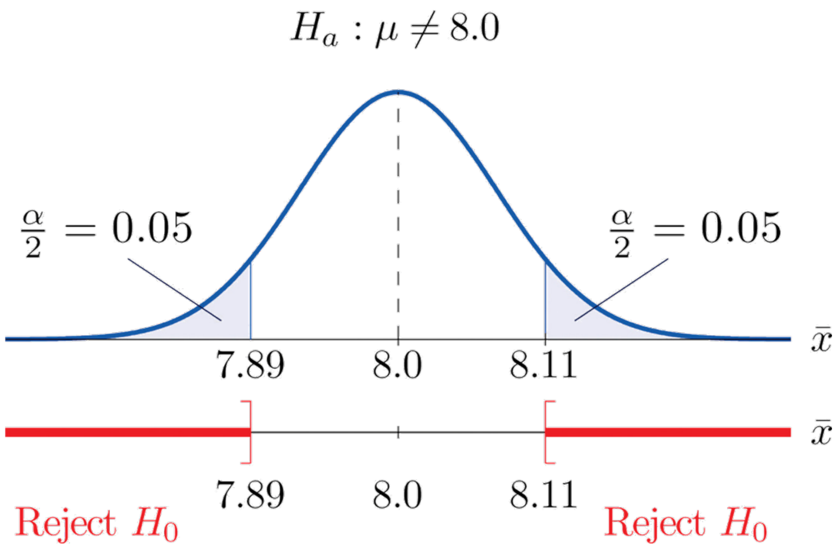
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990

ซึ่งเราจะได้ $z_\alpha = 1.645$ และได้ค่า E เป็น

$$E = z_\alpha \frac{\sigma}{\sqrt{n}} = 1.645 \times \frac{0.15}{\sqrt{5}} = 0.11$$

นั่นคือขอบเขตทั้งสองฝั่งคือ

$$[\mu_0 - E, \mu_0 + E] = [8.00 - 0.11, 8.00 + 0.11] = [7.89, 8.11]$$



ดังนั้นวิธีการตัดสินใจคือเก็บตัวอย่างขนาด 5 แล้วคำนวณสถิติทดสอบ $\bar{x}$ ถ้าค่า $\bar{x}$ ที่ได้อยู่ในบริเวณปฏิเสธ เราจะสามารถปฏิเสธสมมติฐานหลักได้ แต่ถ้าค่า $\bar{x}$ อยู่ในช่วง $[7.89, 8.11]$ เราจะไม่สามารถปฏิเสธสมมติฐานหลักได้

อย่างไรก็ตามเรามักไม่นิยมใช้วิธีนี้ในการทดสอบสมมติฐาน เรามักจะเปรียบเทียบค่าสถิติกับ**การกระจายตัวแบบเกาเซียนมาตรฐาน (standard Gaussian distribution)** แทน ซึ่งคือการกระจายตัวแบบเกาเซียนรอบค่า 0 และมีค่าเบี่ยงเบนมาตรฐานเท่ากับ 1 ซึ่งเราสามารถพิจารณาได้ไม่ยากว่าตัวแปรสุ่ม

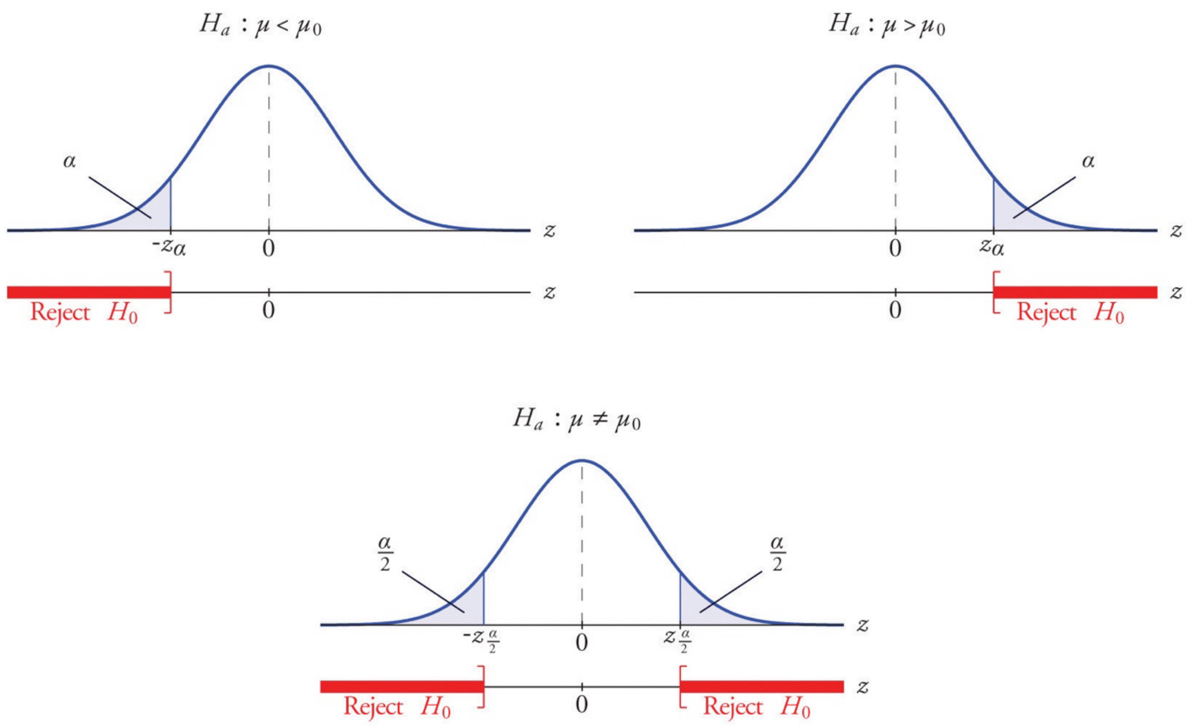
$$\boxed{z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}}$$

มีค่าคาดหวังเป็น 0 และค่าเบี่ยงเบนมาตรฐานเท่ากับ 1 เราจะใช้ตัวแปรนี้เป็นสถิติเพื่อทดสอบสมมติฐานแทน

ในกรณีที่เราไม่ทราบค่าเบี่ยงเบนมาตรฐานของประชากรเราสามารถใช้ค่าเบี่ยงเบนมาตรฐานตัวอย่างเพื่อคำนวณแทนได้ในกรณีที่ตัวอย่างมีขนาดใหญ่ $(n > 30)$ นั่นคือสถิติเพื่อทดสอบสมมติฐานจะคำนวณจาก

$$z = \frac{\bar{x} - \mu_0}{S/\sqrt{n}}$$

จากนั้นให้เปรียบเทียบกับบริเวณปฏิเสธโดยการพิจารณาค่า z_α นั่นคือจากรูป



และตารางด้านบนเพื่อพิจารณาว่าสถิติทดสอบของเราอยู่ในบริเวณปฏิเสธหรือไม่

ตัวอย่าง บริษัทจำหน่ายครีมแห่งหนึ่งเติมครีมทาหน้าลงในขวดโดยเฉลี่ย 8.1 ออนซ์ในหนึ่งขวด โดยมีค่าเบี่ยงเบนมาตรฐาน 0.22 ออนซ์ โดยหน่วยควบคุมมาตรฐาน (quality control, QC) ต้องคอยตรวจสอบไม่ให้เครื่องจักรเติมครีมในขวดน้อยหรือมากเกินไป โดยวิธีการตรวจสอบคือการเลือกตัวอย่าง 30 ตัวอย่างเพื่อนำมาทดสอบสมมติฐาน ในการทดสอบครั้งหนึ่งหน่วยควบคุมมาตรฐานเก็บตัวอย่าง 30 ตัวอย่างและได้ค่าเฉลี่ยของครีมในกระปุกเป็น 8.2 ออนซ์ และมีค่าเบี่ยงเบนมาตรฐานตัวอย่างเป็น 0.25 ออนซ์ จงพิจารณาว่าเครื่องจักรนี้เติมครีมได้มาตรฐานหรือไม่ที่ระดับนัยสำคัญ 1%

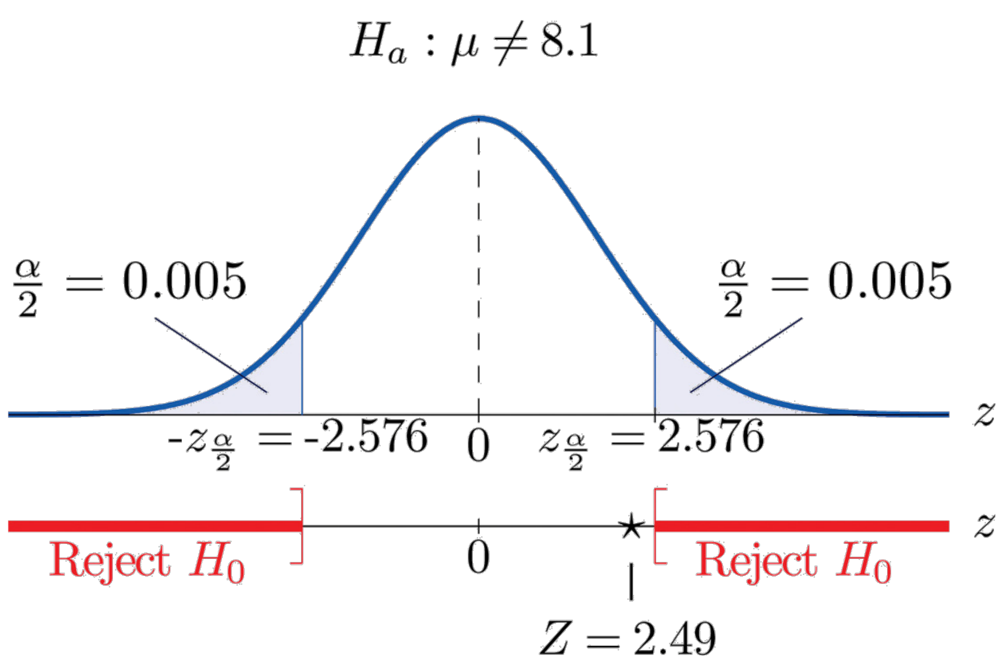
เราจะสร้างสมมติฐานได้เป็น

$$H_0 : \mu = 8.1$$
$$H_a : \mu \neq 8.1, \alpha = 0.01$$

ค่าสถิติทดสอบคือ

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{8.2 - 8.1}{0.22/\sqrt{30}} = 2.490$$

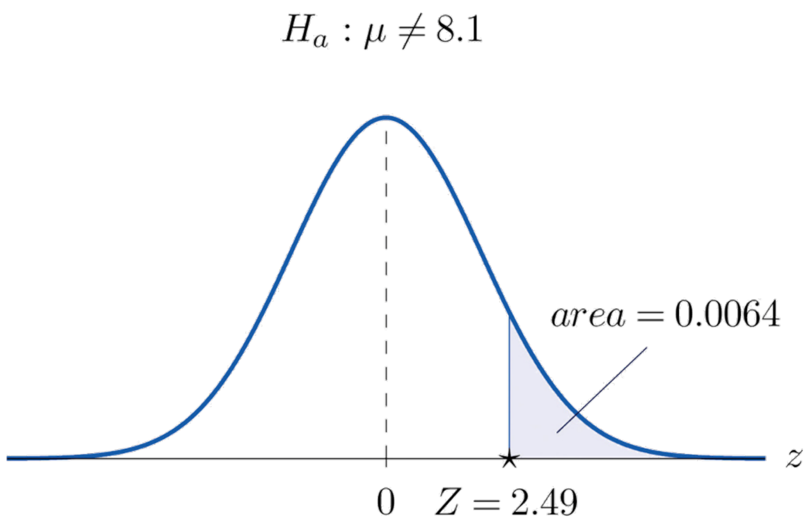
เมื่อพิจารณาค่า $z_{\alpha/2} = z_{0.005}$ จากตารางเราพบว่า $z_{0.005} = 2.576$ ซึ่ง $z < z_{0.005}$ นั่นคือเราจะไม่สามารถปฏิเสธสมมติฐานหลักได้ที่ระดับนัยสำคัญ 1%



อีกวิธีหนึ่งที่นิยมคือ**การทดสอบสมมติฐานด้วย p-value** (probability value) นั่นคือเราจะเปรียบเทียบเป็นความน่าจะเป็นแทน เช่นในตัวอย่างที่แล้วเราคำนวณสถิติทดสอบได้เป็น

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{8.2 - 8.1}{0.22/\sqrt{30}} = 2.490$$

เมื่อเราคำนวณกลับว่ามีความน่าจะเป็นเท่าใดที่ค่า z จะมีค่ามากกว่าที่สังเกตได้นั่นคือจากตารางเราจะพบว่าค่า $z = 2.490$ สอดคล้องกับพื้นที่ด้านขวา 0.0064 และทำให้ค่า $\alpha_{observed} = 2 \times 0.0064 = 0.0128$ นั่นคือ p-value คือ 0.0128 ซึ่งเราจะเห็นว่ามีค่าน้อยกว่าระดับนัยสำคัญที่กำหนดให้ $1.28\% > 1\%$ นั่นคือเรายังไม่สามารถปฏิเสธสมมติฐานหลักได้



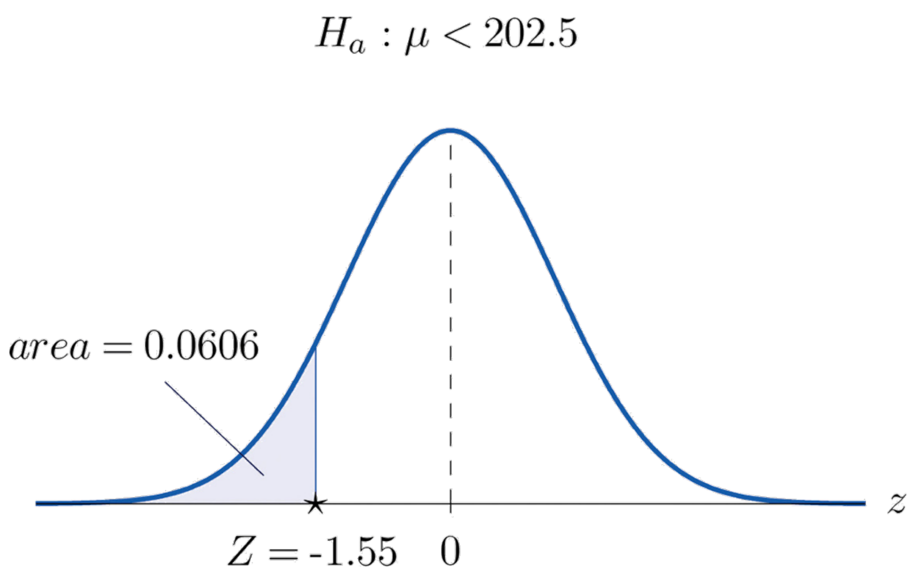
ตัวอย่าง คะแนนรวมของการแข่งบาสเก็ตบอลคือคะแนนของทั้งสองทีมที่ทำได้ในการแข่งหนึ่งครั้ง สำนักข่าวแห่งหนึ่งทำการรายงานว่าคะแนนรวมเฉลี่ยของบาสเก็ตบอล NBA มีค่าเท่ากับ 202.5 อย่างไรก็ตามผู้สนับสนุน NBA หลายคนไม่เชื่อในสำนักข่าวนี้และพวกเขาเชื่อว่าค่าเฉลี่ยต่ำกว่านี้ โดยพวกเขาเลือกตัวอย่างการแข่งขัน 85 ครั้งและได้ค่าเฉลี่ยตัวอย่างเป็น 199.2 และมีค่าเบี่ยงเบนมาตรฐานตัวอย่างเป็น 19.63 จงพิจารณาว่าด้วยระดับความเชื่อมั่น 5% ผู้สนับสนุนมีหลักฐานเพียงพอหรือไม่ที่จะปฏิเสธการรายงานของสำนักข่าวนี้

เราจะสร้างสมมติฐานได้เป็น

ค่าสถิติทดสอบคือ

และหาพื้นที่จากตารางได้เป็น

และผลของการทดสอบคือ



การทดสอบสมมติฐานสำหรับตัวอย่างจำนวนน้อย

สำหรับตัวอย่างที่ผ่านมาเราใช้ central limit theorem เพื่อยืนยันว่าการกระจายตัวของ $\bar{x}$ เป็นแบบเกาเซียนในกรณีที่ขนาดของตัวอย่างมีค่ามากพอ (ข้อตกลงคือมากกว่า 30 ถึงจะใช้ central limit theorem ได้) อย่างไรก็ตามในหลายๆครั้งเราไม่สามารถหาตัวอย่างที่ขนาดมากได้ ในกรณีนี้เราจะแบ่งเป็น 2 กรณี

1. ถ้าเราทราบค่าเบี่ยงเบนมาตรฐานของประชากร σ เราจะให้การแจกแจงของ $\bar{x}$ เป็นเกาเซียนแบบเดิมและใช้สถิติทดสอบเป็น

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

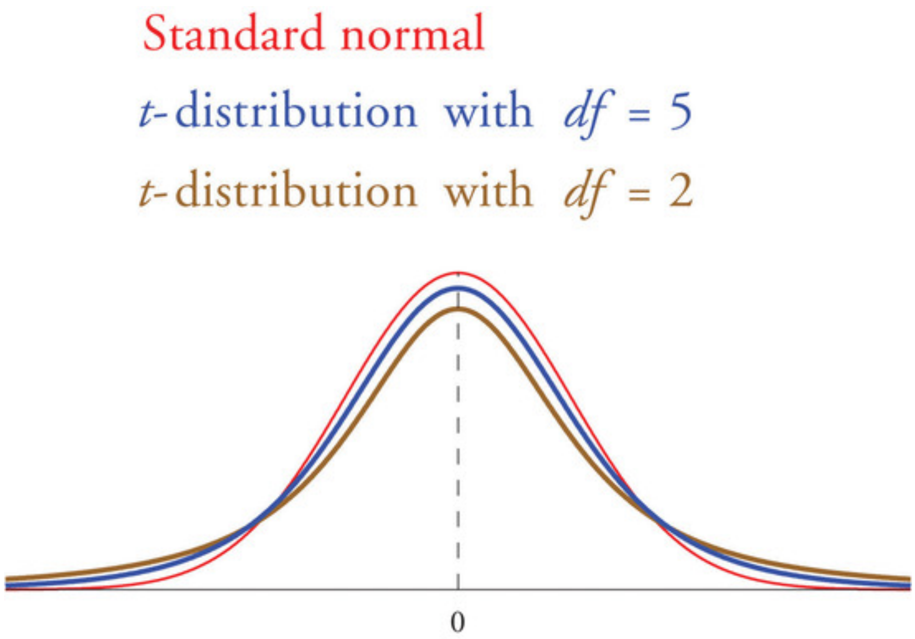
2. ถ้าเราไม่ทราบค่าเบี่ยงเบนมาตรฐานของประชากร เราจะให้การแจกแจงของ $\bar{x}$ เป็นการแจกแจงแบบที (Student's t-distribution) และสถิติทดสอบเป็น

$$t = \frac{\bar{x} - \mu_0}{S/\sqrt{n}}$$

โดยการแจกแจงแบบทีจะมีพารามิเตอร์ที่สำคัญคือองศาอิสระ (degree of freedom)

$$df = \nu = n - 1$$

โดยการแจกแจงทีนี้มีลักษณะคล้ายการแจกแจงแบบเกาเซียนเมื่อจำนวนตัวแปรหรือองศาอิสระมีค่ามาก $n \to \infty$



ν	α						
	0.40	0.30	0.20	0.15	0.10	0.05	0.025
1	0.325	0.727	1.376	1.963	3.078	6.314	12.706
2	0.289	0.617	1.061	1.386	1.886	2.920	4.303
3	0.277	0.584	0.978	1.250	1.638	2.353	3.182
4	0.271	0.569	0.941	1.190	1.533	2.132	2.776
5	0.267	0.559	0.920	1.156	1.476	2.015	2.571
6	0.265	0.553	0.906	1.134	1.440	1.943	2.447
7	0.263	0.549	0.896	1.119	1.415	1.895	2.365
8	0.262	0.546	0.889	1.108	1.397	1.860	2.306
9	0.261	0.543	0.883	1.100	1.383	1.833	2.262
10	0.260	0.542	0.879	1.093	1.372	1.812	2.228
11	0.260	0.540	0.876	1.088	1.363	1.796	2.201
12	0.259	0.539	0.873	1.083	1.356	1.782	2.179
13	0.259	0.537	0.870	1.079	1.350	1.771	2.160
14	0.258	0.537	0.868	1.076	1.345	1.761	2.145
15	0.258	0.536	0.866	1.074	1.341	1.753	2.131
16	0.258	0.535	0.865	1.071	1.337	1.746	2.120
17	0.257	0.534	0.863	1.069	1.333	1.740	2.110
18	0.257	0.534	0.862	1.067	1.330	1.734	2.101
19	0.257	0.533	0.861	1.066	1.328	1.729	2.093
20	0.257	0.533	0.860	1.064	1.325	1.725	2.086
21	0.257	0.532	0.859	1.063	1.323	1.721	2.080
22	0.256	0.532	0.858	1.061	1.321	1.717	2.074
23	0.256	0.532	0.858	1.060	1.319	1.714	2.069
24	0.256	0.531	0.857	1.059	1.318	1.711	2.064
25	0.256	0.531	0.856	1.058	1.316	1.708	2.060
26	0.256	0.531	0.856	1.058	1.315	1.706	2.056
27	0.256	0.531	0.855	1.057	1.314	1.703	2.052
28	0.256	0.530	0.855	1.056	1.313	1.701	2.048
29	0.256	0.530	0.854	1.055	1.311	1.699	2.045
30	0.256	0.530	0.854	1.055	1.310	1.697	2.042
40	0.255	0.529	0.851	1.050	1.303	1.684	2.021
60	0.254	0.527	0.848	1.045	1.296	1.671	2.000
120	0.254	0.526	0.845	1.041	1.289	1.658	1.980
∞	0.253	0.524	0.842	1.036	1.282	1.645	1.960

ν	α						
	0.02	0.015	0.01	0.0075	0.005	0.0025	0.0005
1	15.895	21.205	31.821	42.434	63.657	127.322	636.590
2	4.849	5.643	6.965	8.073	9.925	14.089	31.598
3	3.482	3.896	4.541	5.047	5.841	7.453	12.924
4	2.999	3.298	3.747	4.088	4.604	5.598	8.610
5	2.757	3.003	3.365	3.634	4.032	4.773	6.869
6	2.612	2.829	3.143	3.372	3.707	4.317	5.959
7	2.517	2.715	2.998	3.203	3.499	4.029	5.408
8	2.449	2.634	2.896	3.085	3.355	3.833	5.041
9	2.398	2.574	2.821	2.998	3.250	3.690	4.781
10	2.359	2.527	2.764	2.932	3.169	3.581	4.587
11	2.328	2.491	2.718	2.879	3.106	3.497	4.437
12	2.303	2.461	2.681	2.836	3.055	3.428	4.318
13	2.282	2.436	2.650	2.801	3.012	3.372	4.221
14	2.264	2.415	2.624	2.771	2.977	3.326	4.140
15	2.249	2.397	2.602	2.746	2.947	3.286	4.073
16	2.235	2.382	2.583	2.724	2.921	3.252	4.015
17	2.224	2.368	2.567	2.706	2.898	3.222	3.965
18	2.214	2.356	2.552	2.689	2.878	3.197	3.922
19	2.205	2.346	2.539	2.674	2.861	3.174	3.883
20	2.197	2.336	2.528	2.661	2.845	3.153	3.849
21	2.189	2.328	2.518	2.649	2.831	3.135	3.819
22	2.183	2.320	2.508	2.639	2.819	3.119	3.792
23	2.177	2.313	2.500	2.629	2.807	3.104	3.768
24	2.172	2.307	2.492	2.620	2.797	3.091	3.745
25	2.167	2.301	2.485	2.612	2.787	3.078	3.725
26	2.162	2.296	2.479	2.605	2.779	3.067	3.707
27	2.158	2.291	2.473	2.598	2.771	3.057	3.690
28	2.154	2.286	2.467	2.592	2.763	3.047	3.674
29	2.150	2.282	2.462	2.586	2.756	3.038	3.659
30	2.147	2.278	2.457	2.581	2.750	3.030	3.646
40	2.125	2.250	2.423	2.542	2.704	2.971	3.551
60	2.099	2.223	2.390	2.504	2.660	2.915	3.460
120	2.076	2.196	2.358	2.468	2.617	2.860	3.373
∞	2.054	2.170	2.326	2.432	2.576	2.807	3.291

ตัวอย่าง ไม้ปิงปองตามร้านค้ามาตรฐานถูกขายด้วยราคา 179 บาท ถ้าเราซื้อจากช่องทางออนไลน์จำนวน 5 ร้านค้าแล้วพบว่าราคาขายเป็นดังนี้

$$155, 179, 175, 175, 161$$

จงพิจารณาว่าที่ระดับความเชื่อมั่น 5% การซื้อจากช่องทางออนไลน์ราคาถูกกว่าราคาตามร้านมาตรฐานหรือไม่

ตัวอย่าง ชิ้นส่วนอิเล็กทรอนิกส์ชิ้นหนึ่งมีสองรูเพื่อนำไปต่อกับส่วนอื่นๆของวงจร ในการผลิตถ้าระยะห่างระหว่างสองรูนี้ต้องมีค่าเป็น 0.02 มิลลิเมตร ในระหว่างวันวิศวกรต้องนำตัวอย่างจำนวนเล็กๆจากกระบวนผลิตเพื่อนำไปทดสอบ ถ้าครั้งหนึ่งวิศวกรนำตัวอย่าง 4 ชิ้นและพบว่าระยะห่างมีค่าเป็น

$$0.021, 0.019, 0.023, 0.020$$

จงพิจารณาว่าที่ระดับความเชื่อมั่น 1% กระบวนการผลิตนี้มีปัญหาหรือไม่

แบบฝึกหัด 5

1). การสร้างสมมติฐาน

จงเขียนสมมติฐานหลักและสมมติฐานแย้งของสถานการณ์ต่อไปนี้

1.1) อุณหภูมิเฉลี่ยของเดือนเมษายนของสิบปีที่ผ่านมาคือ 35 องศา เราคิดว่าปีนี้น่าจะร้อนขึ้น

1.2) ราคาห้องพักเฉลี่ยในโรงแรมของประเทศไทยคือ 800 บาท เราคิดว่าราคาห้องพักเฉลี่ยของจังหวัดขอนแก่นสูงกว่านี้

1.3) โรงงานผลิตรถยนต์แจ้งว่าอายุการใช้งานเฉลี่ยของยางคือ 35000 กิโลเมตร เราคิดว่าอายุการใช้งานน้อยกว่านี้

1.4) ซุปเปอร์มาร์เก็ตแห่งหนึ่งต้องการให้การต่อคิวจ่ายเงินไม่นานเกิน 2 นาที ถ้าเราเป็นผู้ตรวจสอบซุปเปอร์มาร์เก็ตแห่งนี้

2). ช่วงปฏิเสธของสมมติฐานหลัก

พิจารณาสมมติฐานหลักและสมมติฐานแย้งต่อไปนี้แล้วหาช่วงปฏิเสธของสมมติฐานหลักเมื่อสมมติให้การแจกแจงของตัวแปรสุ่มเป็นแบบเกาเซียนที่มีค่าเบี่ยงเบนมาตรฐานเป็น σ และทำการทดสอบด้วยตัวอย่างขนาด n

2.1) $H_0 : \mu = 27, H_a : \mu < 27, \alpha = 0.05, \sigma = 5, n = 30$

2.2) $H_0 : \mu = 52, H_a : \mu \neq 52, \alpha = 0.01, \sigma = 10, n = 10$

2.3) $H_0 : \mu = -105, H_a : \mu > -105, \alpha = 0.1, \sigma = 12, n = 20$

2.4) $H_0 : \mu = 78, H_a : \mu \neq 78, \alpha = 0.02, \sigma = 4, n = 5$

3). Z-test

3.1) รัฐบาลของประเทศหนึ่งรายงานอายุขัยเฉลี่ยของประชากรเป็น 66.2 ปี อย่างไรก็ตาม สำนักงานสำมะโนประชากรคิดว่าอายุขัยเฉลี่ยน้อยกว่าที่รายงานจริง จึงได้ทำการสำรวจกลุ่มตัวอย่างจำนวน 30 ตัวอย่างและพบว่าค่าเฉลี่ยตัวอย่างคือ 62.3 ปี และมีค่าเบี่ยงเบนมาตรฐานตัวอย่างเป็น 8.1 ปี จงทดสอบสมมติฐานของสำนักงานสำมะโนประชากรด้วยความเชื่อมั่น 99%

3.2) เมื่อหลายปีก่อนค่าเฉลี่ยของจำนวนสมาชิกในบ้าน 1 หลังคือ 3.14 คน ถ้าเราต้องการสำรวจว่าค่าเฉลี่ยนี้มีการเปลี่ยนแปลงหรือไม่ด้วยความเชื่อมั่น 95% โดยการสุ่มบ้านมาทั้งหมด 75 หลัง และถ้าพบว่าค่าเฉลี่ยตัวอย่างคือ 2.98 คนและมีค่าเบี่ยงเบนมาตรฐานเป็น 0.82 คน ค่าเฉลี่ยนี้มีการเปลี่ยนแปลงหรือไม่

3.3) ปริมาณแคลอรี่ที่นักโภชนาการแนะนำให้ผู้ใหญ่ได้รับต่อวันคือ 2200 แคลอรี่ นักโภชนาการเชื่อว่ามหาวิทยาลัยแห่งหนึ่งมีปริมาณแคลอรี่ที่ได้รับต่อวันน้อยกว่าค่าที่กำหนด จงทดสอบสมมติฐานนี้ที่ความเชื่อมั่น 95% โดยที่ทำการสำรวจตัวอย่างขนาด 36 คน มีค่าเฉลี่ยตัวอย่าง 2150 แคลอรี่ และมีค่าเบี่ยงเบนมาตรฐานตัวอย่างเป็น 203 แคลอรี่

4). p-value

4.1) ซุปเปอร์มาร์เก็ตแห่งหนึ่งต้องการให้การต่อคิวจ่ายเงินไม่นานเกิน 2 นาที ถ้าเราเป็นผู้ตรวจสอบซุปเปอร์มาร์เก็ตแห่งนี้ ต้องการสำรวจตัวอย่างขนาด 30 และได้ค่าเฉลี่ยตัวอย่างเป็น 2.17 นาที และมีค่าเบี่ยงเบนมาตรฐานตัวอย่าง 0.46 นาที จงใช้ข้อมูลเหล่านี้คำนวณค่า p-value พร้อมทั้งเปรียบเทียบกับระดับความเชื่อมั่น 90%

4.2) โรงงานผลิตสินค้าแห่งหนึ่งทราบว่าถ้าเครื่องจักรผลิตสินค้าเฉลี่ยได้มากกว่า 150 ชิ้นต่อชั่วโมงจึงจะได้กำไร ถ้าผู้จัดการได้สั่งซื้อเครื่องจักรใหม่มาจำนวน 36 เครื่องที่ผลิตสินค้าได้แตกต่างกันไปและมีค่าเฉลี่ยตัวอย่างเป็น 160 ชิ้นต่อชั่วโมง โดยมีค่าเบี่ยงเบนมาตรฐานตัวอย่าง 22 ชิ้นต่อชั่วโมง จงหาค่า p-value ของเครื่องจักรชุดนี้ เพื่อตัดสินใจว่าได้กำไรหรือไม่ที่ระดับความเชื่อมัน 99%

5). T-test

5.1) โรงงานผลิตสินค้าแห่งหนึ่งทราบว่าถ้าเครื่องจักรผลิตสินค้าเฉลี่ยได้มากกว่า 150 ชิ้นต่อชั่วโมงจึงจะได้กำไร ถ้าผู้จัดการได้สั่งซื้อเครื่องจักรใหม่มาจำนวน 15 เครื่องที่ผลิตสินค้าได้แตกต่างกันไปและมีค่าเฉลี่ยตัวอย่างเป็น 160 ชิ้นต่อชั่วโมง โดยมีค่าเบี่ยงเบนมาตรฐานตัวอย่าง 10.23 ชิ้นต่อชั่วโมง จงใช้ T-test เพื่อตัดสินใจว่าได้กำไรหรือไม่ที่ระดับความเชื่อมัน 99%

5.2) เหรียญโบราณจำนวน 6 เหรียญถูกค้นพบที่โบราณสถานแห่งหนึ่ง ถ้าน้ำหนักโดยเฉลี่ยของเหรียญโบราณจากบริเวณใกล้เคียงคือ 5.25 กรัม ถ้าเหรียญโบราณหกเหรียญมีน้ำหนักเฉลี่ย 4.73 กรัม และมีค่าเบี่ยงเบนมาตรฐานตัวอย่าง 0.18 กรัม จงทดสอบสมมติฐานว่าเหรียญโบราณชุดนี้มาจากบริเวณนี้หรือไม่ที่ระดับความเชื่อมั่น 99%

5.3) เวลาที่ใช้ในการฟื้นตัวจากการผ่าตัดเข่าโดยเฉลี่ยคือ 123.7 วัน จากประสบการณ์ของนักการภาพบำบัดพบว่าการใช้ยาแก้ปวดบางชนิดทำให้การฟื้นตัวเป็นไปได้ช้ากว่า ถ้าเขาทำการสุ่มตัวอย่างผู้ป่วยจำนวน 7 คนมาและพบว่าเวลาที่ใช้ในการฟื้นตัวคือ

$$128, 135, 121, 142, 126, 151, 123 \text{ วัน}$$

5.3.1) จงทดสอบสมมติฐานนี้ที่ระดับความเชื่อมั่น 90%
5.3.2) จงทดสอบสมมติฐานนี้ที่ระดับความเชื่อมั่น 95%

5.4) เครื่องขายน้ำอัดลมอัตโนมัติแบบหยอดเหรียญถูกออกแบบให้รินน้ำอัดลมในปริมาณเฉลี่ย 16 ออนซ์ต่อถ้วย ในการตรวจสอบการทำงานของเครื่องขายน้ำอัดลมเครื่องหนึ่งโดยการหยอดเหรียญทั้งหมด 9 ครั้งพบว่าปริมาณน้ำอัดลมมีดังนี้

$$15.6, 15.8, 16.2, 16.3, 15.9, 15.5, 15.9, 16.0, 15.8 \text{ ออนซ์}$$

จงทดสอบว่าเครื่องขายน้ำอัดลมนี้ทำงานปรกติหรือไม่ที่ความเชื่อมั่น 95%